

ESSAYS ON THE EFFECTS OF ASYMMETRIC INFORMATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mario Ramirez Basora

August 2012

© 2012 Mario Ramirez Basora
ALL RIGHTS RESERVED

ESSAYS ON THE EFFECTS OF ASYMMETRIC INFORMATION

Mario Ramirez Basora, Ph.D.

Cornell University 2012

It can be easily argued that most, if not all, real economic settings are asymmetric in nature. Particularly, it is often the case that one or several agents possess more or better information than the rest when agreeing upon an economic transaction. Although the information economics revolution of the 1970s laid out the majority of the theoretical foundations, the effects of asymmetric information are subtle and have not been studied in some very interesting contexts, which motivate this thesis.

In the first essay, which is based on joint work with Antonio Bento and Benjamin Ho, we study the problem of an uninformed regulator who wishes to use a voluntary price instrument under varying degrees of uncertainty, specifically in the context of a carbon offset market. In this scenario, a regulator offers private land owners a contract that compensates them for producing carbon offsets while minimizing adverse selection and welfare losses. The model shows that monitoring should decrease as the uncertainty of offset quality decreases, but should increase as uncertainty over agricultural productivity increases. Also, in response to those who argue that the problem of additionality is so large that carbon offsets should not be allowed in carbon regulation, the model quantifies the amount of additionality and finds that even in the case of a regulator with no information, welfare is improved by allowing offset contracts. Finally, the model offers guidance for calculating the optimal offset price as a function of the regulator's information.

The second essay consists of a *cardinal* tournament used by a representative firm to choose its next CEO. Candidates are managers of different types: they are heterogeneous over levels of ability and risk aversion. The managers have private information about their ability. In this context, a two-dimensional solution set of levels of ability and risk aversion corresponding to each possible mean of cash flow realization is identified. Using two different specifications (CARA preferences with normally distributed cash flows, and CRRA preferences with log-normally distributed cash flows), the trade-off between managerial ability and risk aversion is found to be characterized by a concave function. Furthermore, for better levels of technology, the relative importance of risk aversion with respect to ability increases, while for worse levels of technology, the reverse holds.

Finally, in the third essay, using a model based on the optimal consumption and investment models from the operations research literature, I study how the CEO characteristics studied in Chapter 2 impact dividend policy and the long-run evolution of the firm. Specifically, when assuming CRRA preferences and a concave trade-off between ability and risk aversion, I find that the optimal dividend policy of the CEO is non-monotonic with respect to risk aversion. In other words, CEOs with a combination of both high (or low) ability and risk aversion, will pay out lower dividend yields than CEOs with a more balanced combination of ability and risk aversion. Furthermore, firm survival is a function of the dividend yield and is also non-monotonic: while the probability of firm survival converges to either zero or one as risk aversion (and, by extension, ability) converges to either zero or infinity, there exists a range for which lower investment counteracts a potentially higher dividend yield, and the resulting change in the probability of survival is ambiguous.

BIOGRAPHICAL SKETCH

Mario Roberto Ramírez Basora was born on July 27, 1983 in San Pedro de Macorís (Dominican Republic), where he later graduated from Colegio Cristo Rey in 2000. He attended the Pontificia Universidad Católica Madre y Maestra (PUCMM) in Santo Domingo (Dominican Republic), where he received a Bachelor of Science, *cum laude*, in industrial engineering in 2004, as well as a Postgraduate Diploma in business economics in 2005. Before embarking on graduate school, he worked as an industrial engineer at Sara Lee and Johnson & Johnson in his native country. He arrived at Cornell University as a Fulbright scholar in August 2006, where he earned a Master of Arts in economics en route to completing a Doctorate of Philosophy in economics in May 2012.

Para Olga Basora y Dolores Gómez Viuda Basora.

ACKNOWLEDGMENTS

I would like to sincerely thank my special committee members: Professor Antonio Bento, my committee chair, for his indispensable academic guidance and mentorship; Professor Benjamin Ho, for his insightful feedback and advice, as well as his unfaltering support; and Professor David Easley, for his thoughtful comments and suggestions. Professors Bento and Ho have also provided me with the opportunities without which some of my research would not be possible – in particular, Chapter 1 is based on our joint work. Furthermore, Professor Easley, through his course on economics of imperfect information, adeptly and succinctly introduced me to the type of economic research I had vaguely envisioned doing when I arrived at Cornell. Chapters 2 and 3 are based on the research project and ideas from his course.

Many fellow graduate students have made valuable contributions, both academically and socially, which greatly enhanced this dissertation. Of these friends and colleagues, I would like to thank David De Angelis, who helped develop the ideas which evolved into the work presented in Chapters 2 and 3, as well as Carlos Rodríguez Castelán and Joerg Ohmstedt, whose friendships provided fertile intellectual discussions as well as a safe haven from the doldrums of academic research. Most importantly, I would like to thank my fiancéé, Liliana Sousa. Through her thorough readings and critical eye she made my essays significantly better; and, without her unwavering love and support, surviving graduate school would have been infinitely harder.

I want to especially thank my family, to whom I owe everything. I would like to thank my sisters, Gisselle Ramírez and Gina García, for their steadfast and ardent support. Finally, I want to sincerely thank and express my appreci-

ation for my mother, Olga Basora, and my grandmother, Dolores Gómez Viuda Basora. Through them I learned at an early age the importance of knowledge and intellectual curiosity. Through their love, sacrifice and hard work, they instilled in me the principles and values which I have adhered to throughout my personal life, as well as my academic and professional career. Without them, I most surely would not be where I am today.

This research was undertaken with the generous support of several sources of funding. As a Fulbright scholar, the first two years of my doctoral studies were funded by an Alliance for Excellence grant from the Dominican Republic. My remaining years at Cornell were funded by graduate teaching and research assistantships under the guidance of Professors John Abowd, Talia Bar, Antonio Bento, Micky Falkson, Benjamin Ho, Mukul Majumdar and Ted O'Donoghue.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgments	v
Table of Contents	vii
List of Tables	ix
List of Figures	x
 1 Optimal Price Instruments in Voluntary Emission Markets	 1
1.1 Introduction	1
1.2 General Model	7
1.2.1 Firms	7
1.2.2 Regulator	8
1.2.3 Firm Behavior	9
1.2.4 Regulator's Behavior	13
1.3 Linear Case	16
1.3.1 Regulator's Problem	18
1.3.2 Comparative Statics and Optimal Variable Interactions . .	22
1.3.3 Welfare at the Extremes	24
1.3.4 Additionality	26
1.3.5 Division of Firms	29
1.3.6 Summary of Propositions	32
1.4 Conclusion	36
References	37
 2 CEO Selection: A Tournament	 41
2.1 Introduction	41
2.2 General Model	44
2.3 Specifications	49
2.3.1 CARA Preferences/Normal Distribution	51
2.3.2 CRRA Preferences/Log-Normal Distribution	53
2.4 Analysis of Decision Rule (I)	54
2.4.1 CARA Preferences/Normal Distribution	55
2.4.2 CRRA Preferences/Log-Normal Distribution	58
2.5 A Note on Decision Rule (II)	60
2.5.1 N Managers, Uniform Cash Flows	62
2.5.2 N Managers, Normal and Log-Normal Cash Flows	63
2.5.3 $N = 2$ Managers, Normal Cash Flows	65
2.6 Conclusion	66
References	68

3	Dividend Policies and Firm Survival	71
3.1	Introduction	71
3.2	General Model	73
3.2.1	The CEO's Decision Problem	74
3.3	Optimal Dividend and Investment Policies	80
3.4	CRRA Preferences	83
3.4.1	Dividend Policies	86
3.5	A Note on Firm Survival	89
3.6	Conclusion	90
	References	92
A	Appendix to Chapter 1	97
A.1	Adverse Selection Design à la Macho-Stadler and Pérez-Castrillo	97
A.1.1	Firm Behavior	97
A.1.2	Regulator	98
B	Appendix to Chapter 2	105
B.1	Solution for CARA/Normal Distribution Specification	105
B.2	Solution for CRRA/Log-Normal Distribution Specification	106
C	Appendix to Chapter 3	110
C.1	General Model à la Dutta and Radner (1999)	110
C.1.1	Assumptions about Technology	111
C.2	CRRA Preferences/Geometric Brownian Motion Specification	112

LIST OF TABLES

1.1	Types of Firms and Causes of Inefficiency	29
-----	---	----

LIST OF FIGURES

1.1	Historic Volume Growth of the Voluntary Carbon Offset Markets	2
1.2	Historic Value Growth of the Voluntary Carbon Offset Markets .	2
1.3	Common Carbon Offset Types	8
1.4	Real and Estimated Marginal Benefit Curves of Agriculture and Offset Production	11
1.5	3D Continuum and Division of Firms over (θ, β, δ)	21
1.6	2D Continuum and Division of Firms over (θ, β)	22
2.1	Set of Possible Managers and Iso-Mean Curves	48
2.2	Plot of the Function $A(R)$ - CARA/Normal Specification	56
2.3	Plot of the Function $A(R)$ - CRRA/Log-Normal Specification	59
3.1	Heuristic Description of the Model	74
3.2	Dividend Policies and Probability of Survival as a Function of Capital Stock and Severance.	85
3.3	Optimal Dividend Yields as a Function of Risk Aversion for Likely CEOs.	88
C.1	Plot of the Optimal Dividend Policy	115
C.2	Plot of a Sample Path for Capital Stock	116

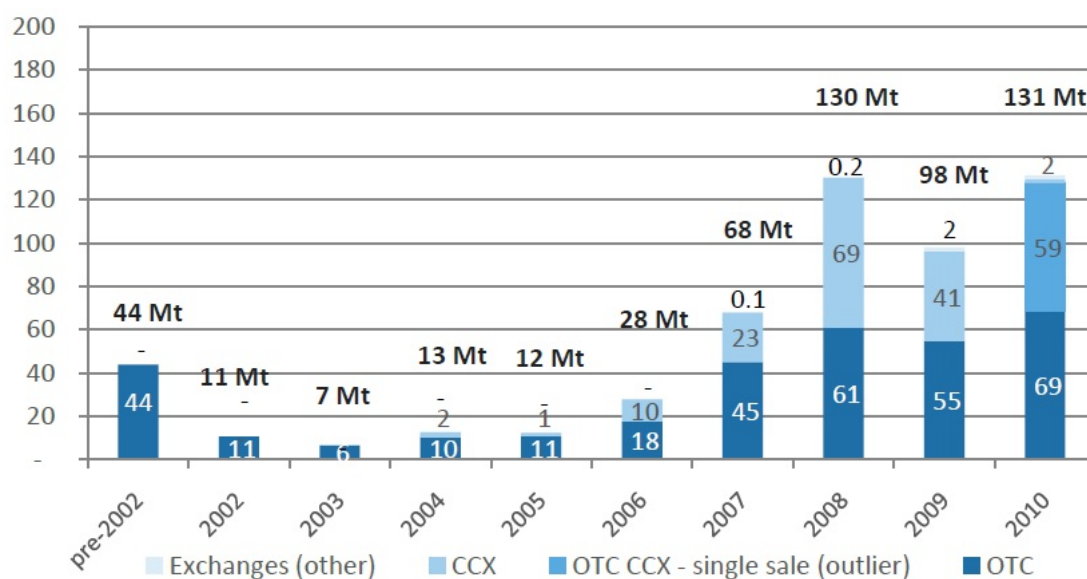
CHAPTER 1

OPTIMAL PRICE INSTRUMENTS IN VOLUNTARY EMISSION MARKETS

1.1 Introduction

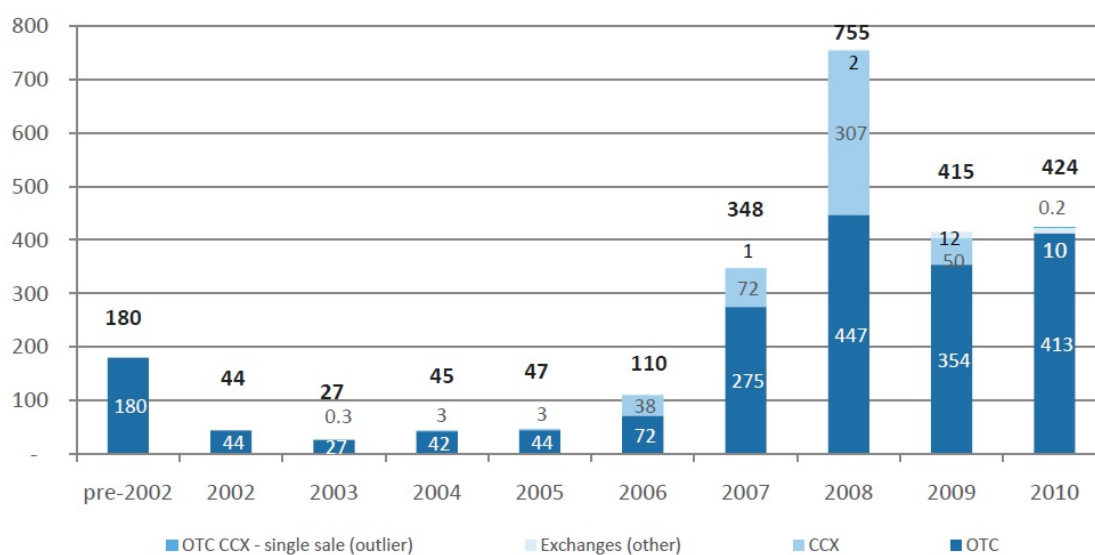
According to the IPCC, agriculture and deforestation together account for a quarter of global anthropogenic greenhouse gas emissions. However, under most proposals to cap emissions (such as Kyoto or the Waxman-Markey Bill in the United States), emissions from sources such as deforestation or agriculture are not capped. Instead, emissions reductions in these sectors are normally incentivized as carbon offset programs where firms receive payments in exchange for agreements to reduce, i.e., offset, their emissions. Furthermore, a small but growing part of global climate mitigation efforts is in voluntary offsets markets which allow individuals or organizations to pay offset originators to make carbon reductions in their name. By observing Figures 1.1 and 1.2, from Ecosystem Marketplace, Bloomberg New Energy Finance, we can appreciate the size and growth of these markets. Before the “Great Recession” of the late 2000s, purchases of voluntary offsets at least doubled in volume annually, and the value of these markets grew at an even faster rate. As the global economy recuperates, that trend, or at least positive growth, is expected to continue for the foreseeable future.

However, there is still a general distrust of whether the greenhouse gas reductions from offset projects are “real” and many have expressed concern that allowing firms in capped sectors to use offsets to reduce their obligations threatens the integrity of cap and trade policies.



Source: Ecosystem Marketplace, Bloomberg New Energy Finance.
 Notes: Based on 153 survey respondents. Annual totals may not equal sum of categories due to rounding.

Figure 1.1: Historic Volume Growth of the Voluntary Carbon Offset Markets



Source: Ecosystem Marketplace, Bloomberg New Energy Finance.
 Notes: Based on 125 survey respondents.

Figure 1.2: Historic Value Growth of the Voluntary Carbon Offset Markets

The U.S. Government Accountability Office (GAO) and the Congressional Research Service (CRS), among others, have identified permanence, leakage and additionality (collectively known as PLA) as the primary concerns that threaten the integrity of carbon offsets:

Permanence: Issues of permanency arise when some carbon reductions (such as afforestation) may be reversed at some point in the future (e.g., if the trees get cut down).

Leakage: The problem of leakage occurs when emissions reductions by one firm or industry indirectly cause emissions from another firm or industry to inefficiently increase.

Additionality: An offset is said to lack additionality if the carbon reduction would have happened anyway, without the payments from the offset purchaser.

Together, these three problems undermine the credibility of offsets markets and highlight the necessity of developing efficient strategies to deal with imperfect information.

The key insight is that the PLA concerns all arise from the inherent difficulty in measuring greenhouse gas emissions from sources like agriculture or deforestation. The difficulties in policy design arise due to asymmetric information, in other words, regulators have less accurate information about emissions than the offset originators. If the uncertainty were symmetric, then mismeasurements should average out; however, asymmetric information introduces the possibility of systemic biases. In this paper, we design a model of asymmetric information in which a regulator wishes to design a price instrument in order

to incentivize the efficient production of carbon offsets by land owners that have private information about their heterogeneous characteristics.¹ The specific uncertainties are over the dimensions of land quality (which determines the costs of emissions reduction) and baseline emissions (which determines the amount of emission reductions for which firms get credit). The model predicts which types of firms are most likely to participate in offset programs, and the extent of mismeasurement due to information asymmetries.

In our model, we specifically focus on the concern of additionality, which relates to the extensive literature on regulation in markets with asymmetric information. In this field, Montero (1999, 2000, 2005 and 2008) is notable among recent work. For example, Montero (1999 and 2000) deals with opt-in programs and the trade-off between efficiency gains of lower cost abatement and excess emissions from adverse selection. He finds that the abatement costs of opt-in facilities are lower to those under the cap, and that the welfare loss from excess emissions is more than made up for in savings on abatement costs. Another novelty of Montero (2008) is the creative use of tools from other disciplines, such as operations research. The application of these tools effectively shift the complications of the the usual models from the price function of offsets or permits to what he calls a payback function (transfers made after the fact to the agents), which lends itself to easier analysis.

Mason and Plantinga (2011), Guiteras, Jack and Oliva (2011) and van Benthem and Kerr (2011) are several recent papers that examine the issue of adverse selection.² Mason and Plantinga, as well as Guiteras, Jack and Oliva, uses

¹In the rest of the paper, we refer to land owners as firms.

²Jack (2011) also conducts a field experiment in Malawi to examine the one-shot allocation problem in environmental markets, and confirms the presence of information asymmetries in these markets and demonstrates that project design affects both the cost effectiveness and the environmental effectiveness of carbon offset projects.

a textbook model of adverse selection primarily to motivate their econometric model: Mason and Plantinga propose a land-use change model, while Guiteras, Jack and Oliva use a modified Becker-DeGroot-Marschak mechanism. However, their theoretical framework does not analyze the core issue that we would like to address: how does welfare change under different levels of information when the regulator uses different policies at her disposal. Van Benthem and Kerr find that increasing the scale of voluntary opt-in programs to reduce emissions in unregulated sectors improves efficiency and reduces transfers to agents. Their approach, however, does not deal with the adverse selection problem directly: they study the trade-off between adverse selection and infra-marginal transfers under different policies without solving for an optimal contract. Overall, none of these papers address our main question: how sensitive welfare loss actually is when the regulator, as opposed to having sophisticated contracts at her disposal, is limited to simple contracts such as a price instrument.

More broadly, this paper relates to the literature on optimal monitoring. For example, Schmutzler and Goulder (1997) consider the trade-off between taxing inputs versus taxing emissions when monitoring is imperfect. While taxation could achieve superior welfare outcomes, our model presumes a political constraint that offset contracts must be voluntary. Millock, Sunding, and Zilberman (2002), like us, allow the amount of monitoring to be a choice variable, but still consider only tax instruments. Stranlund and Dhanda (1999) is perhaps closest in the monitoring literature, considering the optimal amount of monitoring and punishment in a tradable emissions permit system, but their model still exists in an environment of mandatory emissions reductions. Regimes where emissions reductions are driven by offsets, however, differ fundamentally from regimes where emissions reductions are mandatory in that participation is voluntary,

and thus offset contracts must satisfy individual rationality constraints. Our paper focuses on this type of system.

Unlike mandatory regimes where firms are required to meet some level of emissions or face punishment, in voluntary offset regimes, firms are offered a take it or leave it offer of a particular price for the emission reductions they promise to produce. While the price is typically keyed to the prevailing emissions price in the market, regulators effectively vary the price offered to the firm through regulations that discount the amount of credit a firm receives as a function of the quality. For example, firms that sell offsets based on landfill methane effectively get a higher price per ton than firms that sell offsets based on afforestation. This paper thus provides guidance to regulators on how offsets should be discounted and priced.

Finally, let us describe how this paper will develop: we set up the regulator and the agents' maximization problem and study the optimal behavior of the agents. We then study the scenario in which the regulator can only use a voluntary price instrument to regulate the carbon offset market. We also study the welfare loss in this scenario under different levels of information. Here, we find that under full information, the regulator is able to achieve the first-best solution of no welfare loss. Under asymmetric information we find that the regulator is unable to achieve the first-best solution, but we find that an imperfect system is still preferable to not implementing the price mechanism. That is, the benefits of the emissions reductions always outweigh the costs imposed by additivity. Our results are consistent in that if full monitoring is applied under asymmetric information, the regulator is able to generate the entire surplus (it is not optimal due to monitoring costs), and no monitoring is equivalent to ap-

plying the instrument with no information. Finally, we note some preliminary findings of future steps, as well as some possible modeling extensions, such as a simple contract theory model of adverse selection included in Appendix A.

1.2 General Model

1.2.1 Firms

There exists a continuum of profit maximizing firms, which are differentiated by their marginal cost of producing agricultural goods and their baseline emissions. These are modeled as a two-dimensional random variable. θ_i represents the quality of land to produce agricultural goods, while β_i represents baseline emissions that are a byproduct of allocating all available land to agriculture. Hence, different land endowments have differing potentials to either reduce carbon emissions and/or sequester carbon, which we will consider interchangeable and refer to as producing offsets.³ (θ_i, β_i) is defined as follows: $\theta_i \in [\underline{\theta}, \bar{\theta}]$, $\beta_i \in [\underline{\beta}, \bar{\beta}]$, θ_i and β_i are firm-independent and identically distributed over firms according to some joint cumulative distribution $G(\theta_i, \beta_i)$. For now we assume that the marginal densities g_{θ_i} and g_{β_i} exist, while later on we shall make stronger assumptions over the distribution of θ_i and β_i . Firms are endowed with \bar{R} amount of land, which can be devoted to either agriculture or offsets (denoted by R). Production is defined through the production functions $A(\bar{R} - R, \theta_i)$ and $F(R, \beta_i)$, which we assume to be increasing and concave in the first argument (amount of land allocated to production) and increasing in their

³See Figure 1.3 for a better understanding of the common carbon offset projects firms or land owners could possibly engage in.

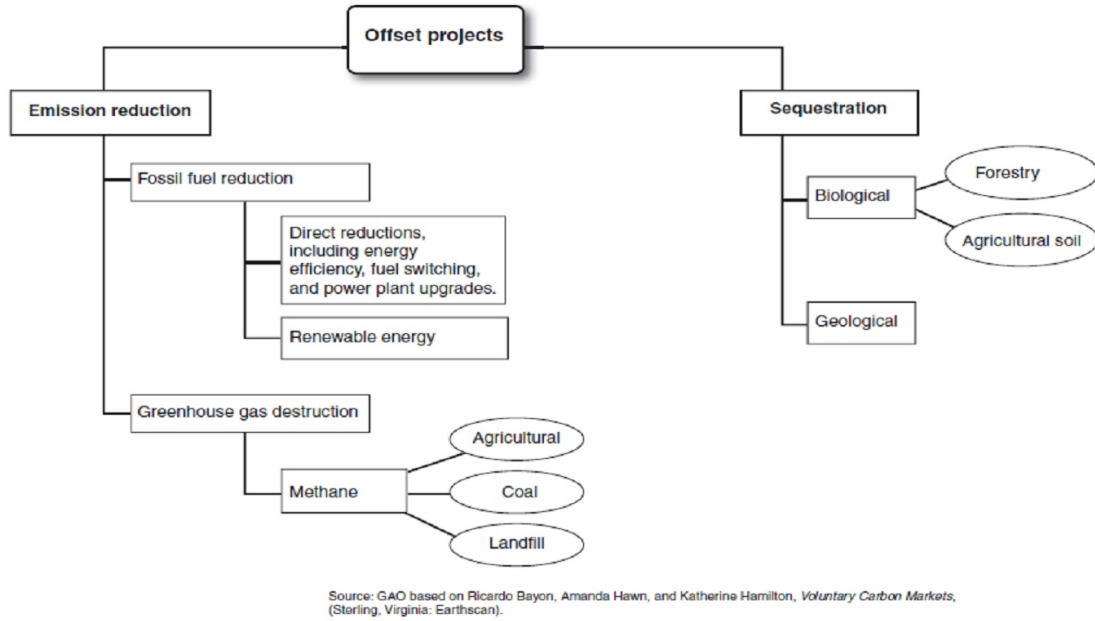


Figure 1.3: Common Carbon Offset Types

respective quality parameter.

1.2.2 Regulator

The regulator wishes to maximize some social welfare function, which we assume is the net sum of total profits from agriculture, the social benefit derived from offsets and the cost of “researching” firms (estimating (θ_i, β_i)). By adding the assumption that the price of the agricultural good is at a fixed equilibrium (which would be valid if, say, the agricultural good is perfectly substitutable with other consumption goods), we believe that this social welfare function accurately captures the welfare components that are subject to change with respect to the implemented policy of incentivizing offset production.

Furthermore, the regulator cannot observe (θ_i, β_i) , so she estimates them:

$(\tilde{\theta}_i, \tilde{\beta}_i)$, where $\tilde{\theta}_i = \theta_i + i(\varepsilon_i, m)$ and $\tilde{\beta}_i = \beta_i + j(\delta_i, m)$, where $\varepsilon_i \in [-\underline{\varepsilon}, \bar{\varepsilon}]$ and $\delta_i \in [-\underline{\delta}, \bar{\delta}]$ are random variables with a cumulative joint distribution $H(\varepsilon_i, \delta_i)$. An implicit assumption is that both $\tilde{\theta}_i$ and $\tilde{\beta}_i$ are elements in the support set of θ_i and β_i . We will make this assumption explicit in the following section, where we study the case of linear production functions and uniformly distributed land quality parameters. The quality of these estimations, as seen in the error terms above, are determined by the regulator's monitoring level, $m \in \mathbb{R}_+$. The cost of monitoring is defined by $C(m)$, which we assume to be increasing and convex. The regulator's problem is to choose the level of monitoring and the price which she will offer for offsets.

For the rest of the paper, without loss of generality, we forgo the i sub-index and observe the behavior of an arbitrary firm and what type of contract it will be offered.

1.2.3 Firm Behavior

Firms maximize profits given their characteristics and the contract they are offered. This contract is a given amount of offsets at a fixed price $p_f(\tilde{\beta}) > 0$ (which will depend on what type of firm the regulator believes it to be). Given a fixed price of the agricultural good, $p_a > 0$, the firm's maximization problem will be:

$$\max_{\bar{R} \in \mathbb{R}_+} \left\{ p_a A(\bar{R} - R, \theta) + p_f(\tilde{\beta}) F(R, \tilde{\beta}) \right\} \quad (1.1)$$

Note that this objective function implies two important points: First, the only choice variable is land; specifically, how to allocate it between projects. Second, the profit the firm generates through producing offsets is actually a function of

the type of land the regulator perceives it to be, rather than its actual type.

The firm's optimal behavior will be defined by the following first order condition:

$$p_a \frac{\partial A(\bar{R} - \hat{R}, \theta)}{\partial R} = p_f(\tilde{\beta}) \frac{\partial F(\hat{R}, \tilde{\beta})}{\partial R} \quad (1.2)$$

where \hat{R} denotes the optimal allocation of land from the firm's perspective. An implicit assumption of this condition is that these two curves intersect at some value of R between 0 and \bar{R} . However, we can easily envision firms with such a high productivity of land, that $p_a \frac{\partial}{\partial R} (A(\bar{R} - R, \theta)) > p_f(\tilde{\beta}) \frac{\partial}{\partial R} (F(R, \tilde{\beta}))$, $\forall R$, and specifically, $p_a \frac{\partial}{\partial R} (A(\bar{R}, \theta)) = p_f(\tilde{\beta}) \frac{\partial}{\partial R} (F(0, \tilde{\beta}))$. In other words, these firms will never produce offsets. For them, this first order condition does not apply, and they simply devote their land to produce agricultural goods. Let us define the fraction of firms that exhibit this characteristic by $\alpha_g > 0$.

Though we will expand more on the regulator's behavior, we should note that for a given estimate $(\tilde{\theta}, \tilde{\beta})$, the regulator believes that there is an optimal allocation of land \tilde{R} that the firm should use for offset production, $\tilde{R} = R(\tilde{\theta}, \tilde{\beta}, p_f)$, which solves the following optimization problem:

$$\max_{\tilde{R} \in \mathbb{R}} \left\{ p_a A(\bar{R} - \tilde{R}, \tilde{\theta}) + p_f(\tilde{\beta}) F(\tilde{R}, \tilde{\beta}) \right\} \quad (1.3)$$

which is solved by the amount of land $R = \tilde{R}$ that satisfies the following first order condition:

$$p_a \frac{\partial A(\bar{R} - \tilde{R}, \tilde{\theta})}{\partial R} = p_f(\tilde{\beta}) \frac{\partial F(\tilde{R}, \tilde{\beta})}{\partial R} \quad (1.4)$$

This amount of land \tilde{R} could be demonstrably greater than, less than, or ambiguous with respect to $\hat{R} = R(\theta, \beta, \tilde{\beta}, p_f)$, depending on the relationship

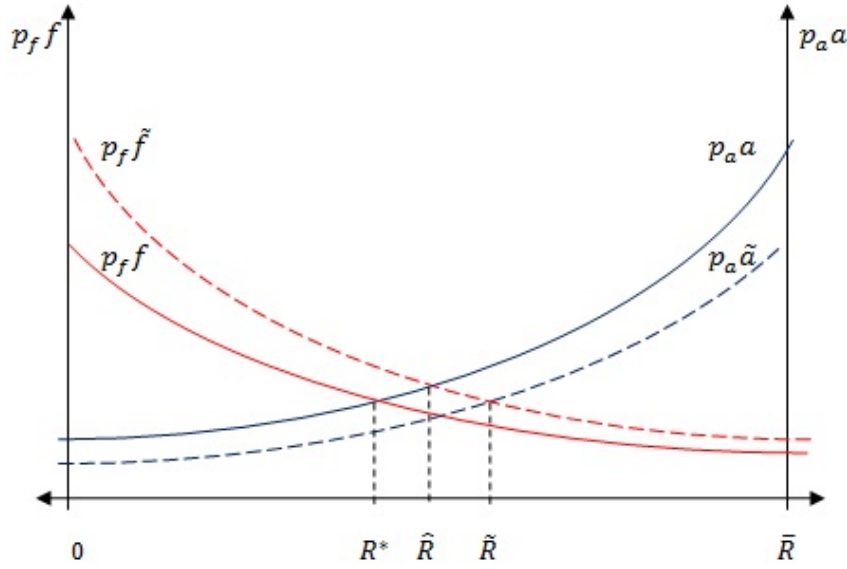


Figure 1.4: Real and Estimated Marginal Benefit Curves of Agriculture and Off-set Production

between the estimates $(\tilde{\theta}, \tilde{\beta})$ and the actual values of (θ, β) . We will also make reference to what we will call the “full information scenario” (where $\theta = \tilde{\theta}$ and $\beta = \tilde{\beta}$), which implies an $R^* = R(\theta, \beta, p_f)$.

Let us study one particular case depicted in Figure 1.4. The firm with the depicted marginal benefit curves is a firm for which the regulator has overestimated β and underestimated θ . This implies that the regulator believes the firm to have a higher offset marginal productivity of land than it actually has:

$$\theta > \tilde{\theta} \implies p_a a = p_a \frac{\partial A(\bar{R} - R, \theta)}{\partial R} > p_a \frac{\partial A(\bar{R} - R, \tilde{\theta})}{\partial R} = p_a \tilde{a}, \quad \forall R \geq 0$$

and a lower agriculture marginal productivity of land:

$$\tilde{\beta} > \beta \implies p_f \tilde{f} = p_f \frac{\partial F(R, \tilde{\beta})}{\partial R} > p_f \frac{\partial F(R, \beta)}{\partial R} = p_f f, \quad \forall R \geq 0$$

In this case, the firm will always allocate more land to offset production than is efficient under perfect information, but less than what the regulator expects the firm to allocate. Mathematically, it chooses \hat{R} such that $R^* < \hat{R} < \tilde{R}$.

Of note in this scenario is that firms will always produce more agriculture and less carbon offsets than what the regulator expects:

$$\begin{aligned} \int_{\hat{R}}^{\bar{R}} p_a \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR &> \int_{\tilde{R}}^{\bar{R}} p_a \frac{\partial A(\bar{R} - R, \tilde{\theta})}{\partial R} dR \\ \iff \int_{\hat{R}}^{\bar{R}} \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR &> \int_{\tilde{R}}^{\bar{R}} \frac{\partial A(\bar{R} - R, \tilde{\theta})}{\partial R} dR \end{aligned}$$

and, since $\tilde{R} > \hat{R}$ and $\tilde{\theta} > \theta$, we have

$$\implies A(\bar{R} - \hat{R}, \theta) > A(\bar{R} - \tilde{R}, \tilde{\theta})$$

Analogously, $F(\hat{R}, \beta) < F(\tilde{R}, \tilde{\beta})$.

However, they are still producing more offsets and less agriculture than what would be expected in the full information case:

$$\begin{aligned} \int_{R^*}^{\bar{R}} p_a \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR &> \int_{\hat{R}}^{\bar{R}} p_a \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR \\ \iff \int_{R^*}^{\bar{R}} \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR &> \int_{\hat{R}}^{\bar{R}} \frac{\partial A(\bar{R} - R, \theta)}{\partial R} dR \end{aligned}$$

and, since $\hat{R} > R^*$

$$\implies A(\bar{R} - R^*, \theta) > A(\bar{R} - \hat{R}, \theta)$$

By the same logic, $F(R^*, \beta) < F(\hat{R}, \beta)$.

For the case in which the regulator overestimates θ and underestimates β , we get the reverse case:

$$\left\{ \begin{array}{l} R^* > \hat{R} > \tilde{R} \\ A(\bar{R} - R^*, \theta) < A(\bar{R} - \hat{R}, \theta) < A(\bar{R} - \tilde{R}, \tilde{\theta}) \\ F(R^*, \beta) > F(\hat{R}, \beta) > F(\tilde{R}, \tilde{\beta}) \end{array} \right.$$

Intuitively, this means that offset production is more land intensive than what the regulator anticipates, yet not as land intensive as they would be in the full information case.

1.2.4 Regulator's Behavior

The regulator will observe noisy signals of (θ, β) , $(\tilde{\theta}, \tilde{\beta})$. Then, she will solve the welfare problem and offer a price $p_f(\tilde{\beta})$ for each ton of carbon emissions the regulator believes would be reduced given the regulator's best guess for β . In a first-best world with perfect information, the price offered would equal the social benefit of reducing each ton of emissions, $p_e > 0$. However, given the uncertainties over land quality, the regulator will optimally distort the price to account for the "quality" of the emissions reduction (which, in fact, is the actual quantity of emissions reductions as opposed to the regulator's noisy expectation).

When solving her optimization problem, the regulator takes into account the ex post realization of how much land was actually allocated, $\hat{R} = R(\theta, \beta, \tilde{\beta}, p_f)$. This decision is then an input into the regulator's social welfare maximization problem, where she must solve for the optimal level of monitoring and the price that she will offer for offsets production. Now, we can proceed to setup and analyze the regulator's problem:

$$\max_{m, p_f} V(\theta, \beta, \tilde{\theta}, \tilde{\beta}, p_f) \quad (1.5)$$

where $V(\theta, \beta, \tilde{\theta}, \tilde{\beta}, p_f) =$

$$\iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[p_a A(\bar{R} - \hat{R}, \theta) + p_e F(\hat{R}, \beta) - C(m) \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) \quad (1.6)$$

$V(\theta, \beta, \tilde{\theta}, \tilde{\beta}, p_f)$ is the total welfare expression given the optimal decision of the firm, \hat{R} . Note that this implies that the regulator has full information about technology: she knows what the production functions look like. The only sources of uncertainty are the specific characteristics of the firm, θ and β .

Given this setup, we can proceed to solve for the regulator's optimal carbon offset price and level of monitoring. The optimality conditions for the regulator (applying Leibniz's rule) are:

$$m^* : \iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[\left(-p_a \frac{\partial A(\bar{R} - \hat{R}, \theta)}{\partial R} \right) \frac{d\hat{R}}{dm} + p_e \frac{\partial F(\hat{R}, \beta)}{\partial R} \right] - C'(m^*) \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) = 0 \quad (1.7)$$

where

$$\frac{d\hat{R}}{dm} = -\frac{\partial \hat{R}}{\partial \tilde{\beta}} \frac{\partial j(\delta, m)}{\partial m}$$

and

$$p_f^* : \iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[\left(-p_a \frac{\partial A(\bar{R} - \hat{R}, \theta)}{\partial R} \right) \frac{\partial \hat{R}}{\partial p_f} + p_e \frac{\partial F(\hat{R}, \beta)}{\partial R} \right] \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) = 0 \quad (1.8)$$

Substituting (1.2) into (1.8), we get:

$$p_f : \iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[\left(-p_f(\tilde{\beta}) \frac{\partial F(\hat{R}, \tilde{\beta})}{\partial R} + p_e \frac{\partial F(\hat{R}, \beta)}{\partial R} \right) \frac{\partial \hat{R}}{\partial p_f} \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) = 0 \quad (1.9)$$

Recall that p_e is the marginal social benefit of each unit of carbon offsets. Note that the price that the regulator will offer is fully dependent on the offset production function, only depending on the agricultural production function in so far as that affects the land allocation decision of the firm. Here we can again see that if $\tilde{\beta} = \beta$, or in other words, $\delta = 0$, we could set $p_e = p_f$ and transform the firm's problem to the regulator's. However, since $\delta \neq 0$ in general, some loss in welfare should be expected to occur. The welfare loss associated to asymmetric information (with respect to the perfect information scenario) should then be

$$\Delta = \hat{V} - V^* = \iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[p_a \left[A(\bar{R} - \hat{R}, \theta) - A(\bar{R} - R^*, \theta) \right] + p_e \left[F(\hat{R}, \beta) - F(R^*, \beta) \right] \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) \quad (1.10)$$

The loss of welfare is, of course, due to the suboptimal allocation of land. This allocation of land contributes to welfare reduction by distorting the level of both agricultural and offsets production. Specifically, welfare is a function of the decrease (or increase) of agricultural output caused by the inefficient allocation of land, and the difference between the actual realization of offsets production and the offsets that would have been produced with the optimal allocation of land.

1.3 Linear Case

In order to allow us to obtain closed form solutions, let us study a particular production function. Namely, let us assume that production is linear in both agriculture and offsets. Specifically,

$$A(\bar{R} - R, \theta) = \theta(\bar{R} - R) \text{ and } F(R, \beta) = \beta R$$

Given this assumption, the firms' maximization problem is now:

$$\max_{\hat{R} \in \mathbb{R}} \left\{ p_a \theta (\bar{R} - R) + p_f (\tilde{\beta}) \tilde{\beta} R \right\} \quad (1.11)$$

The linearity of the production functions implies a corner solution where the firm will only allocate land to agriculture *or* offsets, but not both. The firm's optimal land allocation is given by:

$$\hat{R} = \begin{cases} \bar{R}, & \text{if } p_f (\tilde{\beta}) \tilde{\beta} > p_a \theta \\ R \in [0, \bar{R}], & \text{if } p_f (\tilde{\beta}) \tilde{\beta} = p_a \theta \\ 0, & \text{if } p_f (\tilde{\beta}) \tilde{\beta} < p_a \theta \end{cases} \quad (1.12)$$

Note that the conditions for each value of \hat{R} also define the fractions of firms that will produce either just agriculture or just offsets. Similarly, the allocations of land according to the regulator allowing for asymmetric information \tilde{R} and with perfect information R^* are presented below:

$$\tilde{R} = \begin{cases} \bar{R}, & \text{if } p_f (\tilde{\beta}) \tilde{\beta} > p_a \tilde{\theta} \\ R \in [0, \bar{R}], & \text{if } p_f (\tilde{\beta}) \tilde{\beta} = p_a \tilde{\theta} \\ 0, & \text{if } p_f (\tilde{\beta}) \tilde{\beta} < p_a \tilde{\theta} \end{cases} \quad (1.13)$$

and

$$R^* = \begin{cases} \bar{R}, & \text{if } p_f(\beta) \beta > p_a \theta \\ R \in [0, \bar{R}], & \text{if } p_f(\beta) \beta = p_a \theta \\ 0, & \text{if } p_f(\beta) \beta < p_a \theta \end{cases} \quad (1.14)$$

Now, let us define the measurement error the regulator experiences when trying to estimate $(\tilde{\theta}, \tilde{\beta})$. One of the implications of p_f only being a function of $\tilde{\beta}$ is that it is not necessary to specify a functional form for $\tilde{\theta}$. This is due to how profits are generated: agricultural profit is only function of the firm's real ability to produce agricultural goods. Let us define $\tilde{\beta}$ by imposing some structure on the information production function $j(\cdot)$:

$$\tilde{\beta} \equiv \beta + j(\delta, m) = \beta + \delta, \text{ where } \delta \sim U \left[\frac{\underline{\beta} - \beta}{m+1}, \frac{\bar{\beta} - \beta}{m+1} \right] \quad (1.15)$$

The motivation behind the functional form of the upper and lower bound of the uniform distribution over δ is that we want to, in some sense, properly model monitoring. If the regulator does not perform any monitoring ($m = 0$), we want her to have no good estimate for β : it can be any value between $\underline{\beta}$ and $\bar{\beta}$. On the other hand, as the regulator increases her level of monitoring, the bounds converge towards the real value, β . In other words, as monitoring goes to infinity, the measurement error goes to zero. Mathematically, $\lim_{m \rightarrow \infty} \delta(F_\beta(\beta), m) = 0$. Also, if we assume $\frac{\partial p_f(\beta)}{\partial \beta} > 0$, we can show that:⁴

$$\begin{aligned} \hat{R} = R^* &\iff p_f(\tilde{\beta}) \tilde{\beta} = p_a \theta = p_f(\beta) \beta \\ &\iff \tilde{\beta} = \beta \iff \delta = 0 \text{ or } m = \infty \end{aligned}$$

⁴This condition is proven when solving the regulator's problem

In other words, the regulator can only know the true value of β in a world of asymmetric information by implementing infinite monitoring. This implies that we will always have a welfare loss if there exists some uncertainty over β .

The linearity of the production functions also allows us to partition the space of firms into three different types:

- I. All firms that satisfy $p_f(\tilde{\beta})\tilde{\beta} > p_a\theta$ will only produce offsets.
- II. All firms that satisfy $p_f(\tilde{\beta})\tilde{\beta} < p_a\theta$ will only produce agriculture.
- III. All firms that satisfy $p_f(\tilde{\beta})\tilde{\beta} = p_a\theta$ are indifferent between producing either good.

We will study the firms in categories (I) and (III), with the understanding that all remaining firms in the continuum will belong to category (II).

1.3.1 Regulator's Problem

As before, the regulator first estimates the firm's decision and uses that expected outcome to calculate the welfare maximizing monitoring and offset prices. In the linear production technologies case, this welfare problem is as follows:

$$\max_{p_f, m} V(\theta, \beta, \tilde{\theta}, \tilde{\beta}, p_f) \quad (1.16)$$

where

$$\begin{aligned} V(\theta, \beta, \bar{\theta}, \bar{\beta}, p_f) &= \iint_{\theta, \beta} \left[\iint_{\varepsilon, \delta} \left[p_a \theta (\bar{R} - \hat{R}) + p_e \beta \hat{R} - C(m) \right] dH(\varepsilon, \delta) \right] dG(\theta, \beta) \\ &= \bar{R} \left[\iiint_{\theta, \beta, \delta} p_a \theta dH(\delta) dG(\theta, \beta) + p_e \beta dH(\delta) dG(\theta, \beta) \right] - C(m) \end{aligned}$$

Assuming a uniform distribution over all variables, and assuming that $\bar{\beta} > \frac{p_a}{p_f} \bar{\theta}$ and $\bar{\beta} < \frac{p_a}{p_f} \bar{\theta}$, then we have:⁵

$$= \bar{R} \left[\left(\int_{\bar{\beta}}^{\bar{\theta}} \int_{\frac{\bar{\beta}-\beta}{m+1}}^{\frac{\bar{\theta}-\beta}{m+1}} \left[\int_{\frac{p_f}{p_a}(\beta+\delta)}^{\bar{\theta}} p_a \theta d\theta + \int_{\bar{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} p_e \beta d\theta \right] \frac{m+1}{\Delta\theta(\Delta\beta)^2} d\delta d\beta \right) \right] - C(m)$$

Solving for this integral and simplifying we get:

$$\begin{aligned} &= \frac{\bar{R}}{\Delta\theta} \left[\left[\frac{p_a \bar{\theta}^2}{2} - \frac{p_f^2}{p_a(m+1)} \left(\frac{m^2+1}{2(m+1)} E(\beta^2) + \frac{m}{m+1} E(\beta)^2 \right) \right] + \right. \\ &\quad \left. \left[\frac{p_e p_f}{p_a(m+1)} \left[E(\beta)^2 + m E(\beta^2) \right] - p_e \bar{\theta} E(\beta) \right] \right] - C(m) \end{aligned}$$

From this expression we can derive the new first order conditions:

$$\begin{aligned} p_f^* : \frac{\bar{R}}{\Delta\theta} \left[-\frac{2p_f}{p_a(m+1)} \left(\frac{m^2+1}{2(m+1)} E(\beta^2) + \frac{m}{m+1} E(\beta)^2 \right) \right. \\ \left. + \frac{p_e}{p_a(m+1)} \left[E(\beta)^2 + m E(\beta^2) \right] \right] = 0 \quad (1.17) \end{aligned}$$

$$m^* : \frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{p_a} \text{Var}(\beta) \left[p_e + p_f + m(p_e - p_f) \right] = C'(m) \quad (1.18)$$

⁵Respectively, these assumptions are made to ensure that there exists a positive number of firms for which offset production (or inaction) and agricultural production is always optimal.

From (1.17) we can solve for the optimal offset price, p_f^* :

$$p_f^* = p_e \frac{\left[E(\beta)^2 (m+1)^2 + \text{Var}(\beta) (m^2 + m) \right]}{\left[E(\beta)^2 (m+1)^2 + \text{Var}(\beta) (m^2 + 1) \right]} \quad (1.19)$$

Substituting (1.19) into (1.18) gives us:

$$\implies \frac{\bar{R}}{\Delta\theta} \frac{p_e^3}{p_a} \text{Var}(\beta) \frac{\left[E(\beta^2) + E(\beta)^2 \right] \left[E(\beta)^2 + mE(\beta^2) \right]}{\left[E(\beta^2) (m^2 + 1) + 2E(\beta)^2 m \right]^2} = C'(m) \quad (1.20)$$

Note that the optimal offset price that is offered by the regulator depends solely on the expected value of β , the uncertainty of β (as captured by its variance), and the amount of monitoring that the regulator will choose. Hence, p_f^* is set equal to the marginal social benefit of each unit of carbon offsets adjusted to reflect the regulator's uncertainty over the quality of the land in the production of offsets. Given this solution for p_f^* , we can graph the firms that produce either agricultural goods or carbon offsets on the (θ, β, δ) plane. The defining property for firms which are indifferent between producing either good is $p_f(\beta + \delta) = p_a\theta$. We can gain some intuition by graphically depicting this equilibrium condition. Figure 1.5 shows the parallelepiped which defines the continuum of firms, as well as the plane that divides it into firms that produce agriculture goods or and those that produce carbon offsets. Specifically, the firms below the plane produce agriculture goods while the firms above it produce carbon offsets.

Futhermore, Figure 1.6 shows the cross-section of the (θ, β, δ) body of firms when $\tilde{\beta} = \beta$. For illustrative purposes, we assume $m = 1$. Firms in the area labeled α_o are those that will always choose to produce carbon offsets. For them,

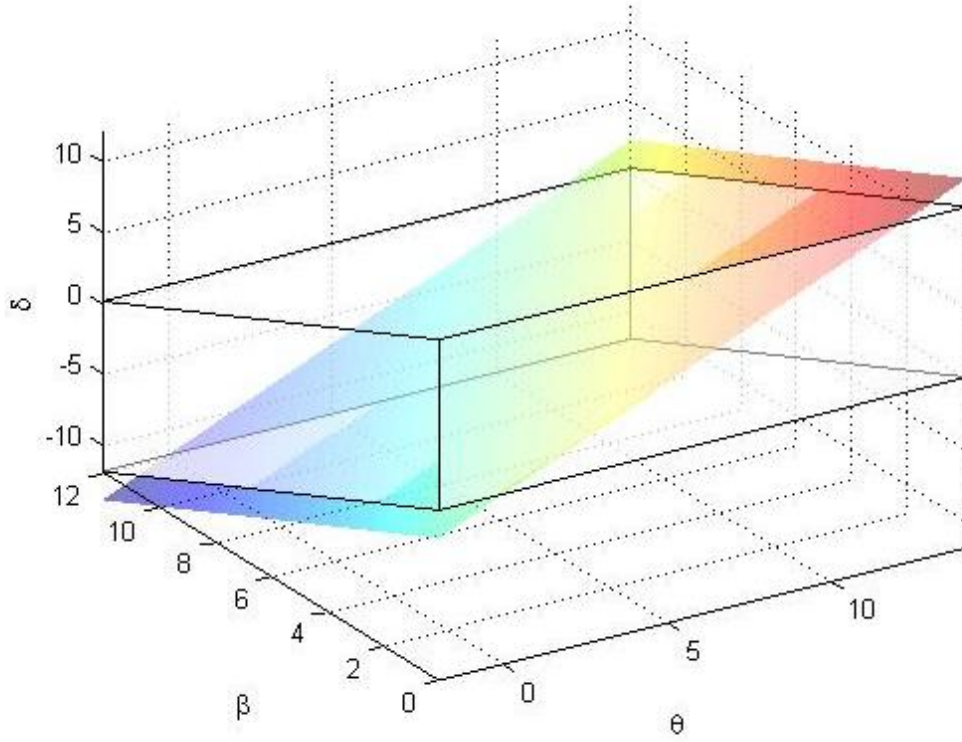


Figure 1.5: 3D Continuum and Division of Firms over (θ, β, δ)

$p_f \tilde{\beta} > p_a \theta$ for all price levels since θ is negative. Without a carbon offset policy, these firms would still have produced offsets since not producing agriculture goods is their optimal choice. Firms in α_2 will also choose to produce carbon offsets, given the p_f being offered. For these firms, a lower offset price would have sufficed to induce them to produce carbon offsets. Firms in α_2 and α_g will choose to continue to produce agriculture goods since the price of offsets being offered by the regulator results in lower returns from offset production than from agriculture. For firms in α_g , this will always be the case – for these firms, $p_f(\tilde{\beta})\tilde{\beta} < p_a \theta$ holds.

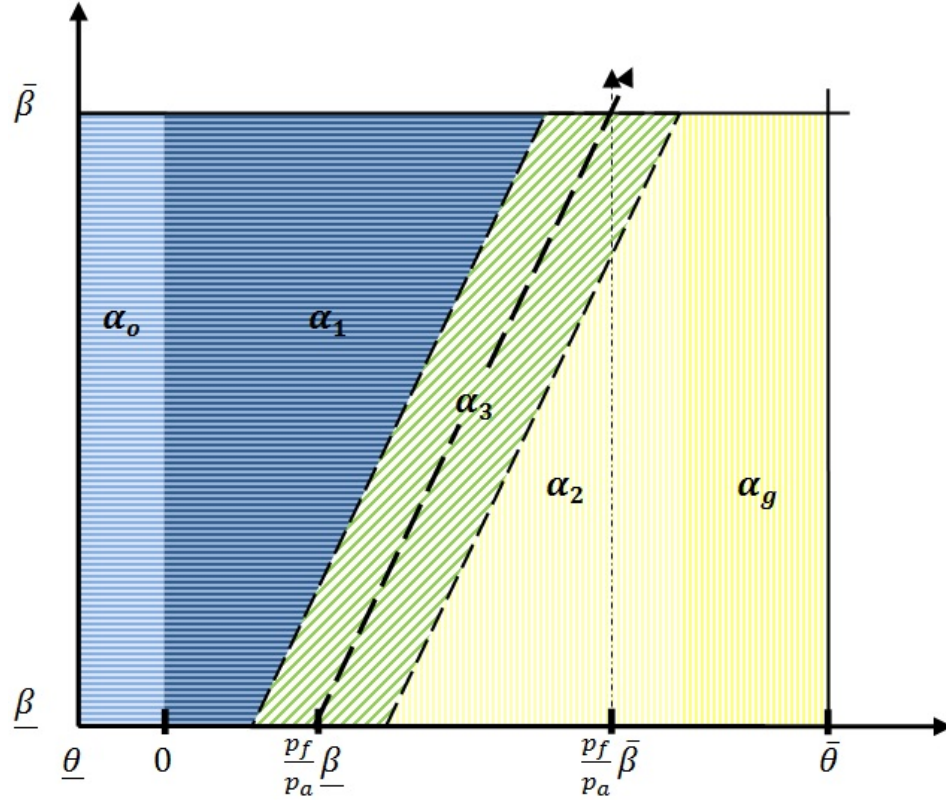


Figure 1.6: 2D Continuum and Division of Firms over (θ, β)

1.3.2 Comparative Statics and Optimal Variable Interactions

The intuition behind the expression for p_f^* , the optimal carbon offset price, is clear when $m \rightarrow \infty$: as monitoring increases and we approach full information, the optimal carbon offset price approaches the marginal benefit of environmental protection ($p_f^* \rightarrow p_e$). At lower levels of monitoring, asymmetric information leads to inefficient pricing schemes. As increases in monitoring lead to better information on the quality of the land, the regulator will pay the farmer exactly the marginal environmental benefit gained from that offset.

Another result that can be derived from the optimality conditions is that the need for monitoring decreases as the difference between $\bar{\beta}$ and $\underline{\beta}$, the variation

in quality of land with respect to production of carbon offsets, becomes smaller.

We can prove that m goes to zero in the limit: if $\bar{\beta} = \underline{\beta} = B$, we have

$$\left[\frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{p_a} \frac{2}{(m+1)^3} \right] * 0 = 0 = C'(m^*)$$

By the assumption that $C(m)$ is increasing and convex, we know that $C'(m^*) = 0 \iff m^* = 0$. The price of offsets at the limit then becomes $p_f^* = p_e \left[E(\beta)^2 / E(\beta^2) \right] = p_e$. As expected, this is the same outcome seen with perfect monitoring, since in both cases, the regulator has perfect information and is able to price the offsets exactly at their environmental value.

On the other hand, as the variation over agricultural soil quality shrinks, the relative value of monitoring increases. To see this, note that as $\Delta\theta \rightarrow 0$, $C'(m^*) \rightarrow \infty \iff m^* = \infty$. Intuitively, as firms become more similar in terms of agricultural quality, the p_f each is willing to accept also becomes more similar. Thus, the regulator will need to choose an offset price more carefully so as to insure that the proportion of firms that accept it will not be excessively high or low. Because of this, monitoring becomes increasingly valuable.

Finally, since $m \geq 0$, we know that $\frac{[E(\beta)^2(m+1)^2 + \text{Var}(\beta)(m^2+m)]}{[E(\beta)^2(m+1)^2 + \text{Var}(\beta)(m^2+1)]} < 1$ if and only if $m < 1$. This implies that the optimal p_f will be below p_e if $m < 1$ and above p_e if $m > 1$. More importantly, the price p_f is increasing over the range $m \in [0, 1 + \sqrt{2} \frac{\sqrt{E(\beta^2) + E(\beta)^2}}{E(\beta^2)}]$, and decreasing over $m > \sqrt{2} \frac{\sqrt{E(\beta^2) + E(\beta)^2}}{E(\beta^2)}$. Intuitively, for low levels of monitoring, the regulator must increase the incentive to firms in order to generate higher participation in the program. As monitoring increases, however, the offset price converges downwards to the social benefit p_e due to the increased quality of information.

1.3.3 Welfare at the Extremes

We now explore the results of this model as the level of information and the price mechanism vary between extremes

No Price Mechanism

With no price mechanism, which necessarily means no information, the resulting welfare is

$$V_0 = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} - p_e \theta E(\beta) \right] \quad (1.21)$$

No Information with Price Mechanism

With no information, but assuming the linear contract studied above, we have a welfare level of

$$V_{NI} = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} + \frac{p_e^2}{2p_a} \frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} - p_e \theta E(\beta) \right] \quad (1.22)$$

Notice that implementing a price mechanism always results in an increase in welfare:

$$\Delta V = V_{NI} - V_0 = \frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{2p_a} \left[\frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} \right] > 0$$

Even though the regulator does not monitor and therefore has no reliable information on the specific values of β , implementing the price mechanism is welfare increasing. This result is driven by the assumption that the regulator knows the distribution of β . However, this is a realistic assumption given that,

though a regulator may not know the exact quality of a given piece of land, she has access to plenty of data to potentially form an informed prior over land quality.

Perfect Information

With perfect information (without the need to monitor), we have

$$V_{PI} = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} + \frac{p_e^2}{2p_a} \left[\text{Var}(\beta) + E(\beta)^2 \right] - p_e \theta E(\beta) \right] \quad (1.23)$$

Hence, the maximum welfare loss attributable to not implementing the price mechanism is

$$\Delta V = V_{PI} - V_0 = \frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{2p_a} \left[\text{Var}(\beta) + E(\beta)^2 \right] > 0$$

Asymmetric Information

We have already computed the level of welfare associated with having asymmetric information – we did this when solving for the regulator’s optimization problem – but it is useful to note the welfare gain by implementing the price mechanism:

$$\begin{aligned} \Delta V = V_M - V_0 &= \frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{p_a} \left[\frac{\left[E(\beta)^2 + m E(\beta^2) \right]^2}{2 \left[(m^2 + 1) E(\beta^2) + 2m E(\beta)^2 \right]} \right] - C(m) \\ &= \frac{\bar{R}}{\Delta\theta} \frac{p_e}{2p_a} \frac{p_f}{m + 1} - C(m) \end{aligned}$$

This allows us make our final observation on the optimal choices of m and p_f : $C(m)$ must be such that $\Delta V(m^*) > 0$, or, $C(m^*) < \frac{\bar{R}}{\Delta\theta} \frac{p_e}{2p_a} \frac{p_f}{m^* + 1}$. In other words, if the cost of monitoring is too high, the regulator will rationally choose not to monitor.

1.3.4 Additionality

Firms which produce offsets are characterized by the condition $p_a\theta < p_f\tilde{\beta}$. However, firms that satisfy the property $\theta < 0$ would have produced offsets anyway. Since the regulator cannot exclude any firm from participating (her only available instrument is to offer an implicit price for offsets where the per unit offset price is the same offered to all firms but the estimation of offsets quantity production is firm specific), we have a number of non-additional offsets. Intuitively, since the agricultural production drops out of the regulator's decision, as seen above, the firm's opportunity cost of entering into an offsets contract with the regulator is not accounted for. Hence, firms whose agricultural production functions are such that they would not produce agricultural goods will always enter into a profitable contract with the regulator and get paid for doing what they would have done absent the carbon offsets program.

To see how this additionality problem is affected by information, let us again examine each scenario of information:

No Price Mechanism

$$F_0 = \bar{R} \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^0 \frac{\beta}{\Delta\beta\Delta\theta} d\theta d\beta = -\bar{R} \frac{\theta}{\Delta\theta} E(\beta) \quad (1.24)$$

F_0 is the number of offsets produced with no price mechanism. Note that this number is positive, since $\underline{\theta} < 0$ by assumption (there is a positive number of firms that have no incentive to farm their land). These are the non-additional offsets.

No Information with Price Mechanism

$$\begin{aligned}
F_{NI} &= \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\beta}-\beta}^{\bar{\beta}-\beta} \int_{\underline{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} \frac{\beta}{(\Delta\beta)^2 \Delta\theta} d\theta d\delta d\beta \\
&= -\frac{\underline{\theta}}{\Delta\theta} E(\beta) + \frac{p_f}{p_a} \frac{E(\beta)^2}{\Delta\theta} = \underbrace{-\bar{R} \frac{\underline{\theta}}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e}{p_a} \frac{\bar{R}}{\Delta\theta} \left[\frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} \right]}_{\text{Additional Offsets}}
\end{aligned} \tag{1.25}$$

where we have plugged in for p_f ($m = 0$). Here we can see that by implementing the program, even with no additional information on land quality, we can increase the number of offsets produced. Since the second term, the additional offsets, is strictly positive for positive prices and nonnegative average emissions reduction capability, any price mechanism results in a number of firms reallocating some land from agricultural production to offsets production.

Perfect Information

$$\begin{aligned}
F_{PI} &= \bar{R} \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^{\frac{p_e}{p_a}(\beta+\delta)} \frac{\beta}{\Delta\beta \Delta\theta} d\theta d\beta = \underbrace{-\bar{R} \frac{\underline{\theta}}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e}{p_a} \frac{\bar{R}}{\Delta\theta} [\text{Var}(\beta) + E(\beta)^2]}_{\text{Additional Offsets}}
\end{aligned} \tag{1.26}$$

Looking at the other information extreme, where the regulator has perfect

information, shows an increase in the number of offsets produced.⁶ Thus, the no information case always results in an inefficiently low production level of offsets.

Asymmetric Information

$$\begin{aligned}
F_M &= \bar{R} \int_{\underline{\beta}}^{\bar{\beta}} \int_{\frac{\beta-\bar{\beta}}{m+1}}^{\frac{\bar{\beta}-\beta}{m+1}} \int_{\underline{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} \beta \frac{(m+1)}{(\Delta\beta)^2 \Delta\theta} d\theta d\delta d\beta \\
&= -\bar{R} \frac{\theta}{\Delta\theta} E(\beta) + \underbrace{\frac{p_f}{p_a} \frac{\bar{R}}{\Delta\theta} \frac{[E(\beta)^2 + mE(\beta^2)]}{m+1}}_{\text{Additional Offsets}} \\
&= \underbrace{-\bar{R} \frac{\theta}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e}{p_a} \frac{\bar{R}}{\Delta\theta} \frac{[E(\beta)^2 + mE(\beta^2)]^2}{(m^2+1)E(\beta^2) + 2mE(\beta)^2}}_{\text{Additional Offsets}} \tag{1.27}
\end{aligned}$$

Under asymmetric information and the inclusion of a price mechanism, the number of additional offsets is a function of how much monitoring the regulator chooses. When $m = 0$, we confirm our result for no information. Also, when $m \rightarrow \infty$, the number of additional offsets converges to $\frac{p_e}{p_a} \frac{1}{\Delta\theta} E(\beta^2)$, which we will see in our next result, matches up with the number of offsets when the regulator has full information. One last note: the number of additional offsets is increasing in m , which also confirms our intuition that as we increase monitoring, the number of offsets increases. This occurs because adverse selection is reduced (firms with low values of β drop out and firms with higher values of β opt in), but, as we shall see in the next segment, this occurs through two mechanisms.

⁶This result follows from $E(\beta^2) = \text{Var}(\beta) + E(\beta)^2$.

Table 1.1: Types of Firms and Causes of Inefficiency

Firms	Action
$p_e\beta < p_a\theta < p_f\tilde{\beta}$	Produce offsets when agriculture would be optimal
$p_f\tilde{\beta} < p_a\theta < p_e\beta$	Produce agriculture when offsets would be optimal
$p_a\theta < p_f\tilde{\beta} < p_e\beta$	Produce offsets, but compensation is less than optimal
$p_a\theta < p_e\beta < p_f\tilde{\beta}$	Produce offsets, but compensation is more than optimal

1.3.5 Division of Firms

Let us now study which and how many firms opt in under each scenario of information quality. We derive closed form expressions for these quantities. The number of firms in each scenario will differ depending on the amount of information the regulator has (in other words, it will depend on m). It is also important to note that under each of these scenarios we have several forms of inefficiency that can occur (see Table 1.1).

No Price Mechanism

Recall that we have assumed that firms are distributed uniformly over the relevant ranges of the (θ, β) space. Hence, we can normalize the number of firms with respect to the total area of firms (which is equal to $\Delta\beta\Delta\theta$). This can perhaps be seen more easily by noting that since we assumed a continuum of firms distributed evenly over a specific area, the integral of the joint cumulative distribution function is equal to 1. After this, we simply multiply by the relevant area. Our normalized number of firms is now:

$$\alpha_o = \Delta\beta\Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^0 \frac{1}{\Delta\beta\Delta\theta} d\theta d\beta = \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^0 1 d\theta d\beta = -\underline{\theta}\Delta\beta \quad (1.28)$$

This is precisely the number we get by observing that the number of firms

that would never produce agriculture even if the policy were not implemented is the rectangle with sides $\Delta\beta$ and $(0 - \underline{\theta}) = \underline{\theta}$.

No Information with Price Mechanism

$$\begin{aligned}\alpha_{NI} &= \Delta\beta\Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\beta}-\beta}^{\bar{\beta}-\beta} \int_{\underline{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} \frac{1}{(\Delta\beta)^2\Delta\theta} d\theta d\delta d\beta \\ &= \Delta\beta \left[-\underline{\theta} + \frac{p_f}{p_a} E(\beta) \right] = \underbrace{-\underline{\theta}\Delta\beta}_{\alpha_o} + \underbrace{\frac{p_e}{p_a}\Delta\beta \frac{E(\beta)^3}{\text{Var}(\beta) + E(\beta)^2}}_{\alpha_1}\end{aligned}\quad (1.29)$$

The area of firms α_1 can be interpreted as a trapezoid with parallel sides of length $\frac{p_f}{p_a}\beta$ and $\frac{p_f}{p_a}\bar{\beta}$, and height $\Delta\beta$. Then, after plugging in for p_f when $m = 0$, we get our final expression for α_1 .

Perfect Information

$$\alpha_{PI} = \Delta\beta\Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^{\frac{p_e}{p_a}\beta} \frac{1}{\Delta\beta\Delta\theta} d\theta d\beta = -\underline{\theta}\Delta\beta + \frac{p_e}{p_a}\Delta\beta E(\beta) \quad (1.30)$$

Under perfect information and with a price mechanism, the share of firms that will enter into carbon offsets contracts will always be larger than in the no information case.⁷ Now we can note the intuitiveness of α_1 : as $\text{Var}(\beta) \rightarrow 0$, $\alpha_{NI} \rightarrow \alpha_{PI}$. This meshes well with our previous result that monitoring being less and less important as the uncertainty over β is reduced.

⁷This is implied by $\text{Var}(\beta) = E(\beta^2) - E(\beta)^2 > 0$.

Asymmetric Information

$$\begin{aligned}
\alpha_M &= \Delta\beta\Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\frac{\underline{\beta}-\beta}{m+1}}^{\frac{\bar{\beta}-\beta}{m+1}} \int_{\underline{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} \frac{(m+1)}{(\Delta\beta)^2\Delta\theta} d\theta d\delta d\beta \\
&= \Delta\beta \left[-\underline{\theta} + \frac{p_f}{p_a} E(\beta) \right] \\
&= \underbrace{-\underline{\theta}\Delta\beta}_{\alpha_o} + \underbrace{\frac{p_e}{p_a}\Delta\beta \frac{(m+1) \left[E(\beta)^2 + mE(\beta^2) \right] E(\beta)}{2mE(\beta)^2 + (m^2+1)E(\beta^2)}}_{\alpha_1+\alpha_3} \tag{1.31}
\end{aligned}$$

The same geometric interpretation made above can be made here: the additional number of firms is just the trapezoid $\alpha_1 + \alpha_3$ seen in Figure 1.6. Note, however, that changes in m lead to changes in p_f . This results in more firms opting in than in the no information scenario. As m increases, p_f will also increase. A higher p_f will make it more desirable for some firms to enter. However, as m increases the range over which β is estimated is reduced. This implies that the total payment to the firm to produce carbon offsets, which is a function of both the offsets price and the regulator's estimation of the quantity of offsets produced by the land, can either increase or decrease.

It turns out that for $m \in \left[0, 1 + \sqrt{2} \frac{\sqrt{E(\beta^2) + E(\beta)^2}}{E(\beta^2)} \right]$, the number of firms that opt in increases: the net effect of increasing the price is that more firms, even firms which are not necessarily good at producing offsets, will opt in. However, for $m > \bar{m} = 1 + \sqrt{2} \frac{\sqrt{E(\beta^2) + E(\beta)^2}}{E(\beta^2)}$, we observe that the number of firms in the program decreases gradually: there is a weeding out process of bad firms, and even though there is a net loss in firms, the fact that the remaining firms are the better ones at producing offsets will imply a greater number of total offsets. Also note that at $m = 1$, we have the same number of firms as when $m = \infty$, though the change in the composition of firms due to monitoring produces differing

levels of total offset production.

To summarize this result, we find adverse selection for all values of m , but, as m increases, this adverse selection converges to zero. To expand on this, for the range $m \in [0, \bar{m}]$ there is a net increase in firms that opt in. We can describe this range as “casting a wider net”, and, as we move along $m \in [\bar{m}, \infty)$, firms that should not have opted in but chose to for small levels of m , now in fact decide against producing offsets.

1.3.6 Summary of Propositions

Proposition 1.1 *Under perfect information, the regulator is able to implement the first-best option of pricing offsets at the price p_e , the marginal social benefit from a unit of carbon offset. This follows from the regulator’s ability to perfectly discriminate. She will not purchase any offsets from firms which fall into the region $p_e \underline{\beta} < p_a \theta$ nor any firm will sell offsets if they are in the region $p_e \bar{\beta} < p_a \theta$.*

Welfare under this regime is

$$V_{PI} = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} + \frac{p_e^2}{2p_a} \left[\text{Var}(\beta) + E(\beta)^2 \right] - p_e \underline{\theta} E(\beta) \right]$$

The number of firms which will opt in will be

$$\alpha_{PI} = \Delta\beta \Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^{\frac{p_e}{p_a} \beta} \frac{1}{\Delta\beta \Delta\theta} d\theta d\beta = -\underline{\theta} \Delta\beta + \frac{p_e}{p_a} \Delta\beta E(\beta)$$

And the number of offset produced is

$$F_{PI} = \underbrace{-\bar{R} \frac{\theta}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e}{p_a} \frac{\bar{R}}{\Delta\theta} [\text{Var}(\beta) + E(\beta)^2]}_{\text{Additional Offsets}}$$

Proposition 1.2 *Under no information, the regulator is only able to implement the worst case scenario of the price instrument. In other words, it is equivalent to implementing the price instrument but without any monitoring. The regulator will offer the price $p_f^* = p_e \left[\frac{E(\beta)^2}{\text{Var}(\beta) + E(\beta)^2} \right] < p_e$.*

This implies that there are firms that would have produced offsets under perfect information, but will not under no information. This is the main source of adverse selection. Welfare under this regime is

$$V_{NI} = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} + \frac{p_e^2}{2p_a} \frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} - p_e \theta E(\beta) \right]$$

The number of firms which will opt in is

$$\alpha_{NI} = -\theta \Delta\beta + \frac{p_e}{p_a} \Delta\beta \frac{E(\beta)^3}{\text{Var}(\beta) + E(\beta)^2}$$

And the number of offsets produced will be

$$\begin{aligned} F_{NI} &= \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\beta}-\beta}^{\bar{\beta}-\beta} \int_{\underline{\theta}}^{\frac{p_f}{p_a}(\beta+\delta)} \frac{\beta}{(\Delta\beta)^2 \Delta\theta} d\theta d\delta d\beta \\ &= \underbrace{-\bar{R} \frac{\theta}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e}{p_a} \frac{\bar{R}}{\Delta\theta} \frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2}}_{\text{Additional Offsets}} \end{aligned}$$

Proposition 1.3 *Under the price instrument, the regulator will implement the following optimal discount:*

$$p_f^* = p_e \frac{\left[E(\beta)^2 (m+1)^2 + \text{Var}(\beta) (m^2 + m) \right]}{\left[E(\beta)^2 (m+1)^2 + \text{Var}(\beta) (m^2 + 1) \right]}$$

and

$$\frac{\bar{R}}{\Delta \theta} \frac{p_e^3}{p_a} \text{Var}(\beta) \frac{\left[E(\beta^2) + E(\beta)^2 \right] \left[E(\beta)^2 + m E(\beta^2) \right]}{\left[E(\beta^2) (m^2 + 1) + 2 E(\beta)^2 m \right]^2} = C'(m)$$

The optimal level of monitoring is $m^* > 0$. The optimal offset price p_f can be either greater or less than p_e .

This, as in the no information case, will produce a sub-optimal allocation, though it is a welfare improvement over the no information case and, as shown below, the case in which no price mechanisms are introduced. Welfare under this scenario is

$$V_M = \frac{\bar{R}}{\Delta \theta} \left[\begin{aligned} & \left[\frac{p_a \bar{\theta}^2}{2} - \frac{p_f^2}{p_a (m+1)} \left(\frac{m^2+1}{2(m+1)} E(\beta^2) + \frac{m}{m+1} E(\beta)^2 \right) \right] \\ & + \left[\frac{p_e p_f}{p_a (m+1)} \left[E(\beta)^2 + m E(\beta^2) \right] - p_e \theta E(\beta) \right] \end{aligned} \right] - C(m^*)$$

Observe that $V_M(m^* = \infty) = V_{PI}$ and $V_M(m^* = 0) = V_{NI}$. From our previous comparison of welfare between no information and perfect information combined with m^* being an interior solution, we know that $V_{PI} > V_M(m^*) > V_{NI}$.

The number of firms that opt in are

$$\alpha_M = -\theta \Delta \beta + \frac{p_e}{p_a} \Delta \beta \frac{(m+1) \left[E(\beta)^2 + m E(\beta^2) \right] E(\beta)}{2m E(\beta)^2 + (m^2 + 1) E(\beta^2)}$$

And the number of offsets produced are

$$F_M = \underbrace{-\bar{R} \frac{\theta}{\Delta\theta} E(\beta)}_{\text{Non Additional Offsets}} + \underbrace{\frac{p_e \bar{R}}{p_a \Delta\theta} \frac{\left[E(\beta)^2 + m E(\beta^2) \right]^2}{(m^2 + 1) E(\beta^2) + 2m E(\beta)^2}}_{\text{Additional Offsets}}$$

Proposition 1.4 *If no offset program is implemented, by construction $p_f = 0$ and $m = 0$.*

Welfare under this scenario is

$$V_0 = \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} - p_e \theta E(\beta) \right]$$

We can compare V_0 to V_{NI} to determine if, with respect to the entire range of firms, it is desirable to implement such a program:

$$\begin{aligned} V_{NI} &= \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} + \frac{p_e^2}{2p_a} \frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} - p_e \theta E(\beta) \right] > \frac{\bar{R}}{\Delta\theta} \left[\frac{p_a \bar{\theta}^2}{2} - p_e \theta E(\beta) \right] = V_0 \\ &\iff \frac{\bar{R}}{\Delta\theta} \frac{p_e^2}{2p_a} \frac{E(\beta)^4}{\text{Var}(\beta) + E(\beta)^2} > 0 \end{aligned}$$

Hence, it is always beneficial as a whole to implement the price instrument, even if the regulator has no additional information besides her prior over distributions. The number of firms that produce offsets are

$$\alpha_o = \Delta\beta \Delta\theta \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^0 \frac{1}{\Delta\beta \Delta\theta} d\theta d\beta = -\theta \Delta\beta$$

And the number of offsets produced are

$$F_0 = \bar{R} \int_{\underline{\beta}}^{\bar{\beta}} \int_{\underline{\theta}}^0 \frac{\beta}{\Delta\beta \Delta\theta} d\theta d\beta = -\bar{R} \frac{\theta}{\Delta\theta} E(\beta)$$

1.4 Conclusion

We use a carbon offset market context to show that for an uninformed regulator who only has a voluntary price instrument at her disposal, she can offer a contract that compensates private agents for producing carbon offsets while reducing adverse selection and welfare losses. Our results hold under varying degrees of uncertainty. The first-best solution is achievable under perfect information or free monitoring. Under asymmetric information and for positive costs of monitoring, we can identify the inefficiencies generated from the additionality problem created by problems of adverse selection. We also show that the net social benefit of an offsets program is always positive. The model created in this paper extends the carbon offsets literature by showing how a voluntary system of carbon offsets, combined with costly monitoring and imperfect information, compares to systems with no offsets pricing mechanism. Though the problem of additionality can only be solved with complete information or monitoring, the results show that the costs associated with additionality do not offset the gains from pursuing a voluntary carbon offset pricing scheme. Importantly, we find that, even in the absence of any information, a pricing mechanism will always be an improvement over no pricing mechanism.

Finally, in Appendix A, we sketch out the general problem of creating an optimal contract using a game theoretic model of adverse selection.⁸ Future steps that can be undertaken in this line of research include moving to a scenario in which the properties of the land are not independent of each other; they are, realistically, correlated. Preliminary results of this extension indicate that this more complex environment should not qualitatively affect our results,

⁸Mason and Plantinga (2011) approach this problem with a very similar framework.

and, intuitively, as the level of correlation goes to 0 we converge to our previous findings (when correlation goes to ± 1 , the model collapses to the perfect information case). Another possible extension of this work would be to construct a more sophisticated framework for the optimal contracting problem in order to model real-world characteristics. For example: incorporating heterogeneous reservation utilities (or equivalently, heterogeneous costs of emissions reduction), as shown in Jullien (2000), would allow for a more careful study of the trade-off the firm faces when deciding land use; and, following the work of Armstrong (1996), as well as Armstrong and Rochet (1999), incorporating multidimensional screening over different parameters of uncertainty (in our case, agricultural and carbon offset land quality) would allow for a richer monitoring environment.

REFERENCES

- [1] Armstrong, M. 1996. "Multiproduct Nonlinear Pricing." *Econometrica* 64 (1): 51–75.
- [2] Armstrong, M. and J.-C. Rochet. 1999. "Multi-Dimensional Screening: A User's Guide." *European Economic Review* 43 (4): 959–979.
- [3] Guiteras, R. P. Jack, B. K. and P. Oliva. 2011. "Additionality in Developing Countries: Optimal Contracts for Avoided Deforestation." *Working Paper*. http://www.webmeets.com/files/papers/aere/2011/231/GuiterasJackOliva_AERE2.pdf (accessed May 9, 2012).
- [4] Jack, B. K. 2011. "Designing Markets for Carbon Offsets: A Field Experiment in Malawi." *Working Paper*. http://sites.tufts.edu/kjack/files/2011/08/Jack_MalawiTrees.pdf (accessed May 9, 2012)
- [5] Jullien, B. 2000. "Participation Constraints in Adverse Selection Models." *Journal of Economic Theory* 93 (1): 1–47.
- [6] Laffont, J.-J. 1989. *The Economics of Uncertainty and Information*. Cambridge, MA: MIT Press.
- [7] Laffont, J.-J. and J. Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.
- [8] Macho-Stadler, I. and Pérez-Castrillo, J. D. 2001. *An Introduction to the Economics of Information: Incentives and Contracts*. New York: Oxford University Press.
- [9] Mason, C. F. and A. J. Plantinga. 2011. "Contracting for Impure Public Goods: Carbon Offsets and Additionality." *NBER Working Paper #16963*.

- [10] Millock, K., D. Sunding and D. Zilberman. 2002. "Regulating Pollution With Endogenous Monitoring." *Journal of Environmental Economics and Management* 44 (2): 221–241.
- [11] Montero, J. P. 1999. "Voluntary Compliance With Market-Based Environmental Policy: Evidence From the US Acid Rain Program." *Journal of Political Economy* 107 (5): 998–1033.
- [12] Montero, J. P. 2000. "Optimal Design of a Phase-in Emissions Trading Program." *Journal of Public Economics* 75 (2): 273–291.
- [13] Montero, J. P. 2005. "Pollution Markets With Imperfectly Observed Emissions." *RAND Journal of Economics* 36 (3): 645–660.
- [14] Montero, J. P. 2008. "A Simple Auction Mechanism for the Optimal Allocation of the Commons." *American Economic Review* 98 (1): 496–518.
- [15] Peters-Stanley, M., K. Hamilton, T. Marcello, and M. Sjardin. 2011. *Back to the Future State of the Voluntary Carbon Markets 2011*. Ecosystem Marketplace & Bloomberg New Energy Finance.
- [16] Salanié, B. 2005. *The Economics of Contracts: A Primer*. 2nd Edition. Cambridge, MA: MIT Press.
- [17] Schmutzler, A. and L. H. Goulder. 1997. "The Choice Between Emission Taxes and Output Taxes Under Imperfect Monitoring." *Journal of Environmental Economics and Management* 32 (1): 51–64.
- [18] Stranlund, J. K. and K. K. Dhanda. 1999. "Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System." *Journal of Environmental Economics and Management* 38 (3): 267–282.

- [19] United States Government Accountability Office. 2008. *The U.S. Voluntary Market Is Growing, but Quality Assurance Poses Challenges for Market Participants*.
- [20] Van Benthem, A. and S. Kerr. 2011. "Bigger Is Better: Avoided Deforestation Offsets in the Face of Adverse Selection." *Stanford PESD Working Paper #102*.

CHAPTER 2

CEO SELECTION: A TOURNAMENT

2.1 Introduction

The strategy employed in choosing the right personnel, particularly for upper management, is one of the most important yet vexing issues faced by firms, impacting corporate policies as well as the long-run evolution and success of the firm. Unobservable heterogeneity over potential candidates affects the decision of who to hire or promote to upper management, and firms would be best served addressing this concern. In this paper, we explore how a firm, through a *cardinal* tournament, chooses its next CEO from a pool of managers that are heterogeneous along two dimensions: ability and risk aversion. These traits of the manager are particularly interesting to consider since the manager's output can rarely be attributed to just ability or risk aversion. To the extent that a firm wishes to promote the most able manager, the firm must contend with the noisy signals generated from both risk aversion and ability. In fact, Goel and Thakor (2007) show that managers who are overconfident skew the odds in their favor and are more likely to be promoted to CEO than rational ones. They also show that this noisiness is not always detrimental: if the CEO is risk averse, this bias in beliefs could even add value to the firm.

This paper broadly relates to the tournament theory literature, of which Lazear and Rosen (1981) is considered its seminal paper. Tournament theory describes scenarios in which the wage difference between agents is not based on their marginal productivity, but on the relative difference, or rank, of the agents. Lazear and Rosen find that if an appropriately sized reward is given to

the winner of the tournament, the agents select the first-best effort level. They also find that if the signal of the agents' skill (output) is noisier, or if there is a larger pool of candidates, in order to induce the first-best effort level the winner's reward must be increased. Lazear and Rosen's work sparked a number of extensions of their general ideas, most notably by Green and Stokey (1983), Nalebuff and Stiglitz (1983) and Rosen (1986). Nalebuff and Stiglitz incorporate risk aversion and allow for penalties as well as rewards. Green and Stokey show the usefulness of tournaments at netting out the common shocks to employee performance. Rosen extends the basic model to a multiple round elimination tournament and shows that, in this context, prizes must be concentrated among a small and select group of agents to keep incentives strong. Finally, Dye (1984), Lazear (1989) and Ramakrishnan and Thakor (1991) discuss the potential pitfalls of a tournament: non-cooperative behavior, such as sabotage or collusion, may counteract the incentive-aligning effect of large spreads in compensation.

Although our model is not a rank-order tournament, as most in the tournament theory literature, we believe it serves our objective, which is to study the impact of ability and risk aversion on manager performance and potential promotion. With this goal in mind, we build a model of a CEO tournament in a representative firm. The firm chooses its next CEO from a pool of n managers. The managers are known to be heterogeneous in level of ability and risk aversion. However, the firm cannot identify these traits. Managers are assigned to individual projects that they can affect through their ability or risk aversion, or, equivalently, they can choose a specific project from a continuum of available projects. Project cash flows are random variables characterized by its two first moments. Managers control the expected return and volatility (or standard variance) of the project's cash flow.

Furthermore, the firm offers the manager a contract that is linear in the project's cash flow as compensation. We assume that this sensitivity of the compensation to performance would solve any moral hazard problems, i.e., managers dedicate their maximal level of effort to the firm. An equivalent interpretation of our setup is that "contracts do not matter", the managers' wealth variations are due to the change in the value of their holdings.¹ Since most CEOs have a large proportion of their wealth exposed to the firm performance, understanding the impact of CEO risk aversion on corporate policies is critical.²

The firm, meanwhile, observes the mean of the project's cash flow realizations but cannot learn the manager's type. We envision the firm selecting the winner through a specific decision rule. We propose several possible decision rules to be the new CEO and proceed to study one of them in depth. Abstracting from luck, or assuming it is randomly distributed, the realized project performance is the result of the CEO's ability and/or her risk strategy.

Moreover, the manager's ability should impact the set of projects available to them, which is defined through the firm's technology level. In other words, managerial ability may be able to augment – or diminish – the firm's given technological capacity. Hence, we analyze how differing technology levels and the manager's traits affect the firm's promotion choice. We examine two different specifications: CARA preferences with normally distributed cash flows and CRRA preferences with log-normally distributed cash flows. We identify a two-dimensional solution set of levels of ability and risk aversion corresponding for each possible level of return, which is illustrated in Figure 2.1. This graph represents combinations of manager's levels of ability and risk aversion explaining

¹Jensen and Murphy (1990) provide empirical evidence in favor of low powered contracts.

²For example, it is forbidden for executives to hedge their position in the firm; hence risk averse managers tend to underinvest and so destroy some value for shareholders.

a given expected cash flow return. We find that the relationship between ability and risk aversion is concave. Furthermore, we observe that, as the level of technology improves, the relative importance of risk aversion with respect to ability increases (and vice versa). This result suggests that industries or firms with different levels of technology will be best served attempting to identify the traits which are best suited for their specific context.

Finally, we include a note at the end of the paper that analyzes a different decision rule, one based solely on observed cash flow. Though we use this note mainly to illustrate a possible direction of future research, we are able derive an expression for the probability of a given manager winning the tournament and being promoted to CEO. This expression is a function of her characteristics, as well as of the characteristics of the rest of the candidate pool. This indicates that, with a more sophisticated model that allows for game theoretic behavior, we could potentially find different optimal managerial decisions, especially if the managers have heterogeneous preferences over being promoted.

2.2 General Model

We assume a pool of n managers, who are rational (i.e., they have correct beliefs) and have heterogeneous preferences and levels of ability. Preferences are characterized by an increasing, concave utility function $u(\cdot)$, with a relative coefficient of risk aversion, R . Ability, A , is defined by the choice set of projects available to a manager. A and R are random and uncorrelated variables.

The cash flow of the project or unit controlled by manager i follows a certain

distribution, D , characterized by its first two moments:

$$X_i \sim D(\mu_i, \sigma_i^2)$$

The contract offered to the manager is linear in the cash flows realized in her business unit: $\tilde{W} = \bar{w} + \beta \tilde{X}_i$, with $\beta \geq 0$.³ We abstract from contract theory by assuming that the manager does not incur disutility by managing her project. Therefore, there is no adverse selection problem in our model. We make this assumption because we are interested in exploring the relationship between ability and risk aversion, and not in the derivation of an optimal contract. We motivate this assumption by noting that, in the corporate world, managers who are in position to be promoted to CEO most likely always exert maximum effort. Hence, we believe that effort will not vary much, if at all, among CEO candidates.

Managers will choose a mean-volatility pair which will determine the distribution function that characterizes their project's cash flow. We can interpret this as being a choice of a specific project from a continuum of available projects with different parameters, or, equivalently, as an assigned project for which, through unobserved effort, she is able to affect distribution of her assigned project's cash flow. The manager i with risk aversion R_i and ability A_i solves the following optimization problem:

$$\max_{(\mu_i, \sigma_i)} E[\pi_i[u(w + \beta X_i) + P_{CEO}^i] + (1 - \pi_i)[u(w + \beta X_i)]] \quad (2.1)$$

$$\text{subject to } \begin{cases} (\mu_i, \sigma_i) \in T(A_i) \\ X_i \sim D(\mu_i, \sigma_i^2) \end{cases} \quad (2.2)$$

³See Holmstrom and Milgrom (1987) for a discussion on linear contracts and the robustness of their optimality.

where π_i is her prior probability of being promoted to the rank of CEO, P_{CEO}^i is the expected discounted utility of a life-time CEO position, and $T(A_i)$ represents the set of possible mean-volatility pairs given the manager's ability, A_i . The technology function, $T(A_i)$, is firm dependent and does not vary across managers. We also assume that managers have no information with respect to the characteristics of each other. Hence, each manager will believe they have a probability $\pi_i = \pi = 1/n$ of being promoted.

How does the firm select a manager? Three different decision rules appear to us most plausible:

- I. The firm observes a string of different cash flows from the different projects in which a manager has invested in. From this information, the firm estimates μ_i for each manager i and selects the manager with the highest μ_i . For example, if the firm (or, equivalently, its shareholders) are risk neutral, we know that given a sufficient amount of capital (or, equivalently, no credit constraints, the firm's expected profit will be maximized by choosing the highest μ_i possible.
- II. The firm observes a string of different cash flows from different projects in which a manager invests in. From this information, the firm estimates (μ_i, σ_i) for each manager i and selects for the manager with the highest mean-volatility (or Sharpe) ratio, $\frac{\mu_i}{\sigma_i}$. For example, if the firm's shareholders have quadratic utility, the Sharpe ratio will maximize the firm's value. Furthermore, Pestien and Sudderth (1988) show that the policy that maximizes the firm's probability of survival is to pick mean and variance to maximize the equivalent mean-variance ratio.
- III. The firm observes a one-shot realization of each manager's cash flow. The

firm selects the manager who has the highest realization. Given the distribution of each manager's cash flow, we are able to derive the ex ante probability of a specific manager having a higher return.

We will study the first and the third decision rules. We choose the first decision rule because it allows for the most straightforward way of analyzing the comparative statics of A_i and R_i with respect to the manager's optimal mean-volatility pair. Note that the second decision rule will lead the shareholders to learn the managers' types and so will suppress the adverse selection problem.⁴ For this very reason, this option is not interesting for us.

Furthermore, we study the third option because, in a sense, it is the most realistic one: observationally, in the business world, we rarely see managers promoted to C-level positions *solely* due to the soundness of their decisions. Instead, these promotions occur due to visible performance, i.e., realized outcomes. In our context, this implies that the manager will not be selected because she chose the best mean-volatility pair, but because she achieved the largest cash flow realization. However, for this case we need to develop an information structure in our model and so it would be technically hard to handle this decision rule.

Since high ability and low risk aversion will constitute desirable traits in our model, we expect a concave relationship between the two, as depicted in Figure 2.1. This relationship implies some equivalence between managers who have both a high level of risk aversion and high level of ability type and managers who have both a low level of risk aversion and low level of ability. We illustrate

⁴Since the manager's optimal expected return and standard deviation will be functions of R_i and A_i , if the firm can learn or accurately estimate these values, and also knows the type of utility function of the manager, then the firm can back out the managers characteristics.

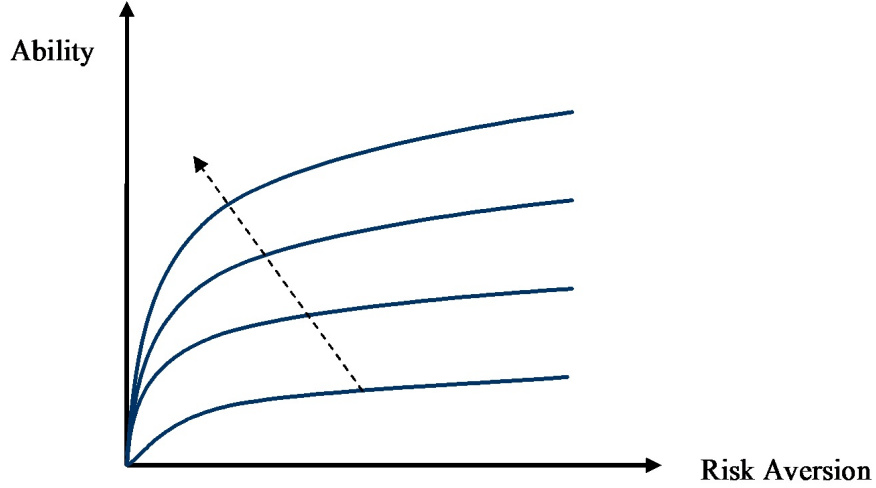


Figure 2.1: Set of Possible Managers and Iso-Mean Curves

a set of manager's coefficients of risk aversion and levels of ability explaining a given mean of cash flows realizations.

Lastly, we observe that, since the utility function is assumed to be increasing, it is necessarily increasing in \bar{w} and in β for a fixed realization of X_i . Hence, we assume $\bar{w} = 0$ and $\beta = 1$ without loss of generality.

Given these assumptions, we can rewrite the manager's objective function as:

$$\begin{aligned}
 & \pi[u(w + \beta X_i) + P_{CEO}^i] + (1 - \pi)[u(w + \beta X_i)] \\
 &= \pi[u(X_i) + P_{CEO}^i] + (1 - \pi)[u(X_i)] \\
 &= u(X_i) + \pi P_{CEO}^i
 \end{aligned} \tag{2.3}$$

Since we assume a fixed probability π , the manager's objective function is a linear transformation of her utility function. This implies that her optimal decisions can be characterized by a standard expected utility maximization problem, without any consideration for promotion. We are satisfied with this setup

because our aim is to explore the relationship and interaction between ability and risk aversion, and not to study scenarios in which the managers have heterogeneous preferences over becoming the CEO. This, along with other ideas, is discussed in the conclusion as a possible extension of our model.

After motivating our assumptions, let us restate and develop the general optimization problem:

$$\max_{(\mu_i, \sigma_i)} \left\{ E[u(X_i)] = \int_{-\infty}^{\infty} u(x) f_{X_i}(x) dx \right\} \quad (2.4)$$

$$\text{subject to } \begin{cases} (\mu_i, \sigma_i) \in T(A_i) \\ X_i \sim D(\mu_i, \sigma_i^2) \end{cases} \quad (2.5)$$

Notice that the objective function, the expected utility of the manager's cash flow, is the expected value of a transformed random variable (where the transformation is the utility function, $u(\cdot)$). Hence, this maximization problem can be solved explicitly for random variables such that, when transformed by concave utility functions, have commonly known (or closed form) probability distribution functions. In the following section, we examine two such scenarios.

2.3 Specifications

In the subsections to follow we examine two different specifications of this optimization problem: (1) the manager has a Constant Absolute Risk Aversion (CARA) utility function and the random variable X_i follows a normal distribution; and, (2) the manager has a Constant Relative Risk Aversion (CRRA) utility function and the random variable X_i follows a log-normal distribution.

We solve for these specifications because they are mathematically tractable, and because they belong to the Hyperbolic Absolute Risk Aversion (HARA) class of utility functions, which is the most general class of utility functions commonly used in economic applications with uncertainty. To illustrate the former, we can transform the expected HARA utility function as follows:

$$\begin{aligned}
E[u(X_i)] &= E \left[\frac{1 - R_i}{R_i} \left(\frac{cX_i}{1 - R_i} + d \right)^{R_i} \right] \\
&= \frac{1 - R_i}{R_i} E \left[\exp \left(\log \left(\frac{cX_i}{1 - R_i} + d \right)^{R_i} \right) \right] \\
&= \frac{(1 - R_i)^{1 - R_i}}{R_i} E \left[\exp \left(R_i \log (cX_i + d(1 - R_i)) \right) \right] \\
&= \frac{(1 - R_i)^{1 - R_i}}{R_i} E [\exp (R_i Z_i)] \\
&= \frac{(1 - R_i)^{1 - R_i}}{R_i} M_Z(R_i)
\end{aligned} \tag{2.6}$$

where

$$M_Z(R_i) = \int_{-\infty}^{\infty} \exp(R_i * z) f_Z(z) dz$$

is the moment generating function of the random variable $Z_i = \log(cX_i + d(1 - R_i)) \sim D(\mu_i, \sigma_i^2)$. This logarithmic relationship is exactly why we proceed to use normal and log-normal probability distributions over cash flows.

We now proceed to (1) solve for the manager's maximization problem for both specifications, and, (2) derive some comparative statics with respect to ability and risk aversion, and examine how a change in these parameters will affect the manager's optimal mean-volatility pair. In Sections 2.4 and 2.5 we will use these results to analyze the relationship between ability and risk aversion with respect to winning the tournament under each of the proposed decision rules.

2.3.1 CARA Preferences/Normal Distribution

Optimal Behavior

In this subsection, we characterize the optimal choice of (μ_i, σ_i) as a function of (R_i, A_i) . For this, we first assume that the utility function is of an exponential form in order to take advantage of its properties (constant relative risk aversion, CARA). Specifically, the utility function is:

$$u_i(X_i) = -\frac{\exp[-R_i X_i]}{R_i} \quad (2.7)$$

Each manager's contract will be defined as a fixed amount plus a fraction of the realized cash flow of her project, which, for the sake of simplicity, we assume are zero and one respectively. We also assume that the manager does not derive more utility from being a CEO (or, equivalently, she receives an additively separable reservation utility normalized to 0). The manager's choice set, which is defined by her technology, is described by the following set: $T(A_i) = \{\forall (\mu_i, \sigma_i) \in \mathbb{R}_{++}^2 : \mu_i \leq A_i (\sigma_i)^\alpha\}$ with $\alpha \in (0, 1)$. For the rest of the paper, we will refer to the parameter α as the technology level. Given this setup, the manager's optimization problem is:

$$\max_{(\mu_i, \sigma_i) \in \mathbb{R}_{++}^2} E \left[-\frac{\exp(-R_i X_i)}{R_i} \right] \quad (2.8)$$

$$\text{subject to } \begin{cases} \mu_i \leq A_i (\sigma_i)^\alpha \\ X_i \sim N(\mu_i, \sigma_i^2) \end{cases} \quad (2.9)$$

Her optimal choice is:⁵

⁵See Appendix B.1. for a description of the optimization problem solution.

$$\mu_i = \left[A_i^2 \left(\frac{\alpha}{R_i} \right)^\alpha \right]^{\frac{1}{2-\alpha}} \quad \text{and} \quad \sigma_i = \left(\frac{\alpha A_i}{R_i} \right)^{\frac{1}{2-\alpha}} \quad (2.10)$$

Comparative Statics

Let us derive some comparative statics on her optimal decision with respect to A_i :

$$\frac{\partial \mu_i}{\partial A_i} = \frac{2}{2-\alpha} \left(\frac{\alpha A_i}{R_i} \right)^{\frac{\alpha}{2-\alpha}} > 0 \quad \text{and} \quad \frac{\partial \sigma_i}{\partial A_i} = \frac{1}{2-\alpha} \left(\frac{\alpha A_i^{\alpha-1}}{R_i} \right)^{\frac{1}{2-\alpha}} > 0 \quad (2.11)$$

Since $\alpha \in (0, 1)$, both derivatives are positive: an increase in the manager's level of ability will induce a riskier choice with a larger expected return as well as higher volatility. Performing the same analysis with respect to R_i yields:

$$\frac{\partial \mu_i}{\partial R_i} = -\frac{\alpha}{2-\alpha} \left[\alpha^\alpha \left(\frac{A_i}{R_i} \right)^2 \right]^{\frac{1}{2-\alpha}} < 0 \quad (2.12)$$

$$\frac{\partial \sigma_i}{\partial R_i} = -\frac{1}{2-\alpha} \left(\frac{\alpha A_i}{R_i^{3-\alpha}} \right)^{\frac{1}{2-\alpha}} < 0$$

Similarly, these results show that if the manager's risk aversion increases, she will choose a safer project with a lower volatility coupled with a lower expected return.

2.3.2 CRRA Preferences/Log-Normal Distribution

Optimal Behavior

Here we follow the same analysis done in the previous subsection but each manager's preferences are now characterized by a CRRA utility function. Specifically,

$$u_i(X_i) = \frac{X_i^{1-R_i}}{1-R_i} \quad (2.13)$$

The rest of the assumptions defined in the previous section remain unchanged. The manager's optimization problem is the following:

$$\max_{(\mu_i, \sigma_i)} E \left[\frac{X_i^{1-R_i}}{1-R_i} \right] \quad (2.14)$$

$$\text{subject to } \begin{cases} \mu_i \leq A_i (\sigma_i)^\alpha \\ \log(X_i) \sim N(\mu_z, \sigma_z^2) \end{cases} \quad (2.15)$$

Hence, her optimal choice is:⁶

$$\mu_i = \left[\frac{A_i^{\frac{2}{\alpha}}}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{\alpha}{2(1-\alpha)}} \quad \text{and} \quad \sigma_i = \left[\frac{A_i^2}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{1}{2(1-\alpha)}} \quad (2.16)$$

Comparative Statics

When differentiating the optimal choices of the manager with respect to A_i and R_i , respectively, we get:

⁶See Appendix B.2. for a description of the optimization problem solution.

$$\frac{\partial \mu_i}{\partial A_i} = \frac{1}{1-\alpha} \left[\frac{A_i^2}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{\alpha}{2(1-\alpha)}} > 0 \quad (2.17)$$

$$\frac{\partial \sigma_i}{\partial A_i} = \frac{1}{1-\alpha} \left[\frac{A_i^{\frac{2}{\alpha}}}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{1}{2(1-\alpha)}} > 0$$

As with the CARA case, both derivatives are positive: an increase in the manager's level of ability will induce a riskier choice with a larger expected return. When increasing her level of risk aversion, R_i , we observe that her optimal choices of μ_i and σ_i will decrease:

$$\frac{\partial \mu_i}{\partial R_i} = -\frac{1}{2} \left[\frac{A_i^2}{\left(\frac{R_i}{\alpha} - R_i - 1\right)^{2-\alpha}} \right]^{\frac{1}{2(1-\alpha)}} < 0 \quad (2.18)$$

$$\frac{\partial \sigma_i}{\partial R_i} = -\frac{\alpha}{2} \left[\frac{A_i^2}{\left(\frac{R_i}{\alpha} - R_i - 1\right)^{3-2\alpha}} \right]^{\frac{1}{2(1-\alpha)}} < 0$$

Since both derivatives are negative, we again have the intuitive result that an increase in the manager's risk aversion will imply a safer choice with a smaller expected return.

2.4 Analysis of Decision Rule (I)

Following the first decision rule described in Section 2.2, managers are selected based on the expected return of their project of choice. This decision rule will

imply a purely deterministic choice by the firm, since it will not depend on an observed realization of a random variable. The firm will simply rank the apparent (or estimated) values of μ_i and select the largest one. The ex post probability of a manager being promoted is either 1 or 0 (where ex post means after the manager's maximization problem is solved). However, we can expand on the relationship between both parameters with respect to the optimal μ_i of the winner of the tournament. We do so for our usual specifications, CARA preferences/Normal distribution and CRRA preferences/Log-normal distribution.

2.4.1 CARA Preferences/Normal Distribution

We now want to study the relative importance of ability versus risk aversion, specifically for different levels of technology. One possible hypothesis is the common perception that ability should be more important when facing a low level of technology, while risk aversion becomes more important when facing a high level of technology. Intuitively, this hypothesis is rooted in the idea that at low levels of technology, the manager must have a high level of ability for the firm to generate profit, while for high levels of technology the firm is better off having a sufficiently low level of risk aversion to take advantage of the high profit projects at its disposal. To study this relationship, let us define $B \equiv \max \{\mu_i\}$ to be the highest observed choice from the pool of managers. For our current specification, the relationship between ability and risk aversion is as follows:

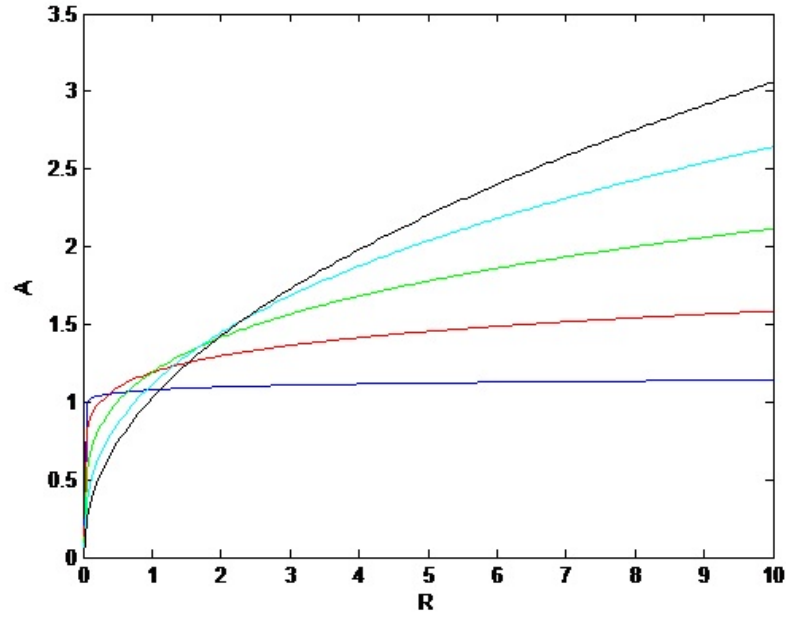


Figure 2.2: Plot of the Function $A(R)$ - CARA/Normal Specification

$$\begin{aligned}
 \max \{\mu_i\} &= B = (A_i)^{\frac{2}{2-\alpha}} \left(\frac{\alpha}{R_i} \right)^{\frac{\alpha}{2-\alpha}} \\
 \implies A_i(R_i) &= B^{\frac{2-\alpha}{2}} \left(\frac{R_i}{\alpha} \right)^{\frac{\alpha}{2}} \\
 \implies A_i(R_i) &\propto \left(\frac{R_i}{\alpha} \right)^{\frac{\alpha}{2}}, \alpha \in (0, 1)
 \end{aligned}$$

Differentiating $A_i(R_i)$, we get:

$$\begin{aligned}
 \frac{\partial A_i(R_i)}{\partial R_i} &= \frac{1}{2} B^{\frac{2-\alpha}{2}} \left(\frac{R_i}{\alpha} \right)^{\frac{\alpha}{2}-1} > 0 \\
 \frac{\partial^2 A_i(R_i)}{\partial R_i^2} &= \frac{1}{4} \frac{\alpha-2}{\alpha} B^{\frac{2-\alpha}{2}} \left(\frac{R_i}{\alpha} \right)^{\frac{\alpha}{2}-2} < 0
 \end{aligned} \tag{2.19}$$

Figure 2.2 provides a plot in (R, A) space of several iso-mean curves for different values of the technological parameter α and for the CARA specification. The different values of α are: 0.05 (in blue), 0.25 (in red), 0.5 (in green), 0.75 (in cyan), and 0.95 (in black).

We see that $A_i(R_i)$ is concave with respect to R_i . This implies that there exists a cutoff point at which ability and risk aversion have the same impact on managerial decisions, which will be dependent on the given technology parameter α . Specifically, at $\frac{\partial A_i(R_i)}{\partial R_i} = 1$, we have $R_i = \alpha 2^{\frac{2}{\alpha-2}}$. Now, we can examine the relative importance of ability and risk aversion.

The condition $R_i = \alpha 2^{\frac{2}{\alpha-2}}$ implies the existence of a set of values of R_i , i.e., $R_i \in \left[0, \alpha 2^{\frac{2}{\alpha-2}}\right)$ for which a manager's risk aversion will have a larger impact on mean return than her level of ability, as well as the existence of the set $R_i \in \left(\alpha 2^{\frac{2}{\alpha-2}}, +\infty\right)$ where a manager's ability will have a larger impact on mean return than her risk aversion. When technology improves ($\alpha \rightarrow 1$), and we restrict ourselves to the range of ability-risk aversion pairs for which risk aversion is relatively more important $\left(\frac{\partial A_i(R_i)}{\partial R_i} > 1\right)$, we find that the level of risk aversion satisfies $R_i < \lim_{\alpha \rightarrow 1} \left\{\alpha \cdot 2^{\frac{2}{\alpha-2}}\right\} = \frac{1}{4}$. Similarly, as $\alpha \rightarrow 0$ and $\frac{\partial A_i(R_i)}{\partial R_i} > 1$, we find that $R_i < \lim_{\alpha \rightarrow 0} \left\{\alpha \cdot 2^{\frac{2}{\alpha-2}}\right\} = 0$.

These results indicate that the set of values for which risk aversion is relatively more important than ability expands as $\alpha \rightarrow 1$ and is equal to $\left[0, \frac{1}{4}\right)$ at the limit. This set reduces to a null set when $\alpha \rightarrow 0$. This matches our ex ante hypothesis of ability being more important for low levels of technology, while risk aversion plays a larger role when the firm has better technology. Conversely, for any level of α , if $R_i > \frac{1}{4}$, the manager's level of ability will always have a larger impact on mean return than her level of risk aversion.

2.4.2 CRRA Preferences/Log-Normal Distribution

Again, let us fix the highest observed choice from the pool of managers as $B = \max \{\mu_i\}$, in order to examine the relationship between ability and risk aversion:

$$\begin{aligned} \max \{\mu_i\} = B &= \left[\frac{A_i^{\frac{2}{\alpha}}}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{\alpha}{2(1-\alpha)}} \\ \implies A_i(R_i) &= B^{1-\alpha} \left[R_i \left(\frac{1}{\alpha} - 1 \right) - 1 \right]^{\frac{\alpha}{2}} \\ A_i(R_i) &\propto \left[R_i \left(\frac{1}{\alpha} - 1 \right) - 1 \right]^{\frac{\alpha}{2}}, \alpha \in (0, 1) \wedge R_i > \frac{\alpha}{1-\alpha} \end{aligned} \quad (2.20)$$

$$\begin{aligned} \frac{\partial A_i(R_i)}{\partial R_i} &= \frac{1}{2} (1-\alpha) \left[R_i \left(\frac{1}{\alpha} - 1 \right) - 1 \right]^{\frac{\alpha}{2}-1} > 0 \\ \frac{\partial^2 A_i(R_i)}{\partial R_i^2} &= \frac{1}{4} (\alpha-2) (1-\alpha) \left[R_i \left(\frac{1}{\alpha} - 1 \right) - 1 \right]^{\frac{\alpha}{2}-2} < 0 \end{aligned}$$

Figure 2.3 provides a plot in (R, A) space of several iso-mean curves for different values of the technological parameter α and for the CRRA specification. As before, the different values of α are: 0.05 (in blue), 0.25 (in red), 0.5 (in green), 0.75 (in cyan), and 0.95 (in black). We can see that $A_i(R_i)$ is concave with respect to R_i . Furthermore, at $\frac{\partial A_i(R_i)}{\partial R_i} = 1$ there exists a cutoff point at which ability and risk aversion have the same impact on managerial decisions. Specifically, $\frac{\partial A_i(R_i)}{\partial R_i} = 1$ holds if and only if $R_i = \left(\frac{\alpha}{1-\alpha} \right) \left[1 + \left(\frac{1-\alpha}{2} \right)^{\frac{2}{2-\alpha}} \right]$. Now, we can examine the relative importance of ability and risk aversion:

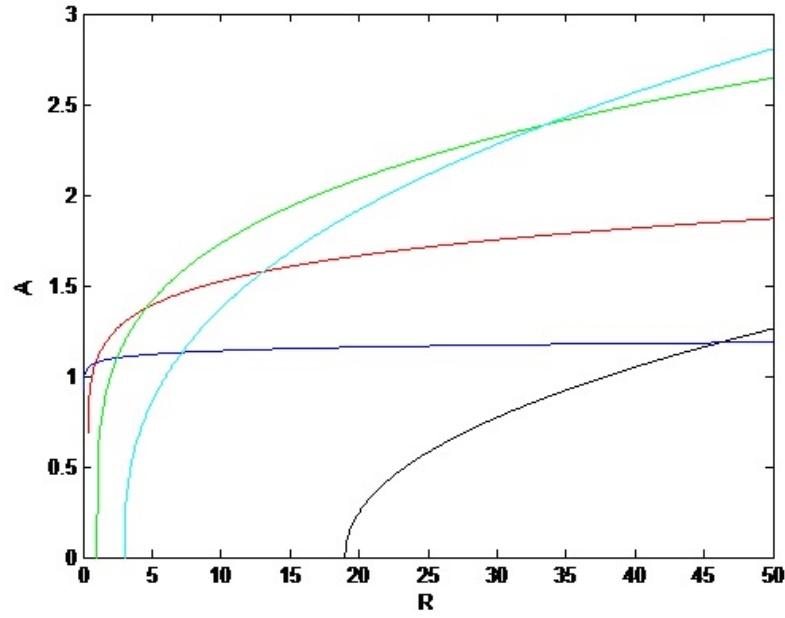


Figure 2.3: Plot of the Function $A(R)$ - CRRA/Log-Normal Specification

$$\text{As } \alpha \rightarrow 1, \frac{\partial A_i(R_i)}{\partial R_i} \geq 1 \iff R_i < \lim_{\alpha \rightarrow 1} \left\{ \left(\frac{\alpha}{1-\alpha} \right) \left[1 + \left(\frac{1-\alpha}{2} \right)^{\frac{2}{2-\alpha}} \right] \right\} = +\infty.$$

$$\text{Similarly, as } \alpha \rightarrow 0, \frac{\partial A_i(R_i)}{\partial R_i} \geq 1 \iff R_i < \lim_{\alpha \rightarrow 0} \left\{ \left(\frac{\alpha}{1-\alpha} \right) \left[1 + \left(\frac{1-\alpha}{2} \right)^{\frac{2}{2-\alpha}} \right] \right\} = 0.$$

These results, together with the constraint $R_i > \frac{\alpha}{1-\alpha}$ which assures a solution to the maximization problem, imply that there exists a set of values of R_i , i.e., $R_i \in \left[\frac{\alpha}{1-\alpha}, \left(\frac{\alpha}{1-\alpha} \right) \left[1 + \left(\frac{1-\alpha}{2} \right)^{\frac{2}{2-\alpha}} \right] \right)$ for which a manager's risk aversion will have a larger impact on mean return than her level of ability. This set is contracting as $\alpha \rightarrow 0$ and as $\alpha \rightarrow 1$ and is equal to a null set at both limits. Finally, another result of interest is that as $\alpha \rightarrow 0$, ability level becomes more important relative to risk aversion.⁷

⁷As $\alpha \rightarrow 1$, the utility maximization problem does not have a solution in the limit.

2.5 A Note on Decision Rule (II)

As described in Section 2.2, the second decision rule chosen to select the future CEO consists simply to select the manager whose project choices (determined through her optimal mean-volatility pair) yield the highest observed cash flow realization. Let us proceed to formulate and derive the probability of a given manager i being promoted.

From our general assumptions, we have that all random variables are independent. However, from our results in the previous sections, we know they are not identically distributed. The key is to notice that, if a manager is promoted to CEO, it will be because:

$$\tilde{X}_i > \max_{j \in \{1, \dots, n\}} \{\tilde{X}_j\}, \quad j \neq i \quad (2.21)$$

Define $X_{-i} \equiv \max_{j \in \{1, \dots, n\}} \{X_j\}$, $j \neq i$ For the general case ($n > 0$), and for fixed probability distribution parameters (i.e., after the managers make their optimal choices), manager i will be selected as CEO with a probability:

$$\begin{aligned}
\hat{\pi}_i &= P(X_i > X_{-i}) = 1 - P(X_i < X_{-i}) = 1 - P(X_i - X_{-i} < 0) \\
&= 1 - F_{X_i - X_{-i}}(0) \\
&= 1 - \int \int_{x \leq y} f_{X_i}(x) f_{X_{-i}}(y) dx dy \\
&= 1 - \int_{-\infty}^{\infty} f_{X_i}(x) dx \int_x^{\infty} f_{X_{-i}}(y) dy \\
&= 1 - \int_{-\infty}^{\infty} \{f_{X_i}(x) dx \cdot [1 - F_{X_{-i}}(x)]\} \\
&= 1 - \left[\int_{-\infty}^{\infty} f_{X_i}(x) dx - \int_{-\infty}^{\infty} F_{X_{-i}}(x) f_{X_i}(x) dx \right] \\
&= 1 - \left[1 - \int_{-\infty}^{\infty} F_{X_{-i}}(x) f_{X_i}(x) dx \right] \\
&= \int_{-\infty}^{\infty} F_{X_{-i}}(x) f_{X_i}(x) dx
\end{aligned}$$

Since we have independence, we can rewrite this as:

$$\hat{\pi}_i = \int_{-\infty}^{\infty} \prod_{j \neq i}^{n-1} F_{X_j}(x) f_{X_i}(x) dx \tag{2.22}$$

This expression has an intuitive interpretation: since the managers' cash flows are independent of each other, the probability of manager i being selected as the new CEO is equivalent to multiplying the probability that manager i achieves some cash flow a by the probability that the rest of the candidate pool all achieve cash flows less than a . We integrate over the support of manager i 's distribution function to account for all possible values of a .

In the following subsections we will discuss two examples: (1) n agents with uniformly distributed cash flows; and, (2) $n \geq 2$ managers with either normal or log-normally distributed cash flows. We examine the scenario in example (1) mainly to illustrate the inner workings of the probability formula, while example (2) serves as a setup of the general case, as well as an analysis of how ability

and risk aversion affect the probability of winning the tournament of a given manager for $n = 2$.⁸

2.5.1 N Managers, Uniform Cash Flows

To properly analyze the manager's optimization problem with uniformly distributed cash flows, we would need to modify the technology set and the choice variables of the firm (since the parameters that define a uniform distribution are its bounds, rather than the mean and variance). Even though we do not make this explicit modeling change, we can still use the uniform distribution to illustrate the intuition of (2.13). We assume that the probability distributions have a support with the same lower bound a but may have distinct upper bounds b_i . We show that, without loss of generality, if manager i has b_i such that $b_i \geq b_j$, $\forall j \neq i$, her probability of winning is $\hat{\pi}_{in} \geq 1/n$:

$$\begin{aligned}
\hat{\pi}_i &= \int_a^{b_i} \prod_{j \neq i}^{n-1} F_{X_j}(x) f_{X_i}(x) dx \\
&= \int_a^{b_i} \prod_{j \neq i}^{n-1} \left(\frac{x-a}{b_j-a} \right) \frac{1}{b_i-a} dx = \frac{1}{b_i-a} \prod_{j \neq i}^{n-1} \left(\frac{1}{b_j-a} \right) \int_a^{b_i} (x-a)^{n-1} dx \\
&= \frac{1}{b_i-a} \sum_{j \neq i}^{n-1} \prod_{j \neq i}^{n-1} \left(\frac{1}{b_j-a} \right) \int_a^{b_j} (x-a)^{n-1} dx \\
&> \frac{1}{b_i-a} \sum_1^{n-1} \prod_{1i}^{n-1} \left(\frac{1}{b_i-a} \right) \int_a^{b_i} (x-a)^{n-1} dx = \prod_{j \neq i}^{n-1} \left(\frac{1}{b_i-a} \right) \frac{(b_i-a)^n}{n} = \frac{1}{n}
\end{aligned}$$

We can also see that if $b_i = b$, $\forall i$, $\hat{\pi}_i = 1/n$. Hence, if the manager had a mechanism through which she could increase her upper bound on cash flow

⁸The expression for $\hat{\pi}$, while seemingly straightforward, quickly becomes intractable for values of $n > 2$ and non identically distributed variables and can only be handled through numerical simulations.

b_i , such as a higher level of ability, she would increase her probability of being promoted to CEO.

2.5.2 N Managers, Normal and Log-Normal Cash Flows

Normal Cash Flows

If the cash flows from the managers' projects are described by a normal random variable, we can compute the following probability that manager i will be promoted:

$$\begin{aligned}\hat{\pi}_i &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} F_{X_j}(x) f_{X_n}(x) dx \\ &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu_j}{\sqrt{2\sigma_j^2}} \right) \right] \frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(x - \mu_n^2)}{2\sigma_n^2} \right) dx\end{aligned}\quad (2.23)$$

Comparative Statics

If managers have CARA utility functions, we can use their optimal choices to generate some comparative statics. By differentiating with respect to A_i and R_i , and using the results derived in Subsection 2.3.1: $\frac{d\mu_i}{dA_i} > 0$, $\frac{d\mu_i}{dR_i} < 0$, $\frac{d\sigma_i}{dA_i} > 0$ and $\frac{d\sigma_i}{dR_i} < 0$ of the CARA/Normal distribution specification, we derive the following natural results:

As the level of ability increases, the resulting change in a manager's probability of winning the tournament is also positive:

$$\frac{\partial \hat{\pi}_i}{\partial A_i} \propto \frac{\partial}{\partial A_i} \left(\frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(x - \mu_n^2)}{2\sigma_n^2} \right) \right) > 0 \quad (2.24)$$

If the manager's risk aversion increases, her probability of winning the tournament will decrease:

$$\frac{\partial \hat{\pi}_i}{\partial R_i} \propto \frac{\partial}{\partial R_i} \left(\frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(x - \mu_n^2)}{2\sigma_n^2} \right) \right) < 0 \quad (2.25)$$

Log-Normal Cash Flows

In the same fashion, if the cash flows from the managers' projects are described by a log-normal random variable, the probability that manager i will be promoted will be:

$$\begin{aligned} \hat{\pi}_i &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} F_{X_j}(x) f_{X_n}(x) dx \\ &= \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\ln(x) - \mu_j}{\sqrt{2\sigma_j^2}} \right) \right] \frac{1}{x\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(\ln(x) - \mu_n^2)}{2\sigma_n^2} \right) dx \end{aligned} \quad (2.26)$$

Comparative Statics

Analogously, if we assume all managers have CRRA utility functions, we can use the comparative statics derived in Subsection 2.3.2: $\frac{d\mu_i}{dA_i} > 0$, $\frac{d\mu_i}{dR_i} < 0$, $\frac{d\sigma_i}{dA_i} > 0$ and $\frac{d\sigma_i}{dR_i} < 0$ of the CRRA/Log-normal distribution specification. Given these results, we have that the probability of manager i winning the tournament increases as A_i grows:

$$\frac{\partial \hat{\pi}_i}{\partial A_i} \propto \frac{\partial}{\partial A_i} \left(\frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(\ln(x) - \mu_n^2)}{2\sigma_n^2} \right) \right) > 0 \quad (2.27)$$

while her probability of winning the tournament decreases as her risk aversion R_i increases:

$$\frac{\partial \hat{\pi}_i}{\partial R_i} \propto \frac{\partial}{\partial R_i} \left(\frac{1}{\sigma_n \sqrt{2\pi}} \exp \left(-\frac{(\ln(x) - \mu_n^2)}{2\sigma_n^2} \right) \right) < 0 \quad (2.28)$$

For both cases, we can also state that if cash flows are i.i.d. random variables, we have $\hat{\pi}_i = \hat{\pi} = 1/n, \forall i$.

2.5.3 $N = 2$ Managers, Normal Cash Flows

To elaborate on the roles that ability and risk aversion play in the probability of a manager being promoted, let us restrict ourselves to the case $n = 2$: we have two managers vying for the position of CEO. In this scenario, the probability of, say, manager 1 being promoted is:

$$P(X_1 > X_2) = P(X_2 - X_1 < 0) = F_X(0) \quad (2.29)$$

where $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ are the random variables that describe the cash flow of their respective projects. Since both random variables are normally distributed, $X = X_2 - X_1$ also follows a normal distribution: $X_2 - X_1 \sim N(\mu_2 - \mu_1, \sigma_1^2 + \sigma_2^2)$. Hence, we can now write manager 1's probability of winning as:

$$F_X(0) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{-\mu_2 + \mu_1}{\sqrt{2(\sigma_1^2 + \sigma_2^2)}} \right) \right] \quad (2.30)$$

Note that, since the error function is positive valued for positive reals, $F_X(0) > 1/2$ if $\mu_1 > \mu_2$. Hence, manager 1 will be the tournament favorite if:

$$\begin{aligned}\mu_1 > \mu_2 &\iff \left[A_1^2 \left(\frac{\alpha}{R_1} \right)^\alpha \right]^{\frac{1}{2-\alpha}} > \left[A_2^2 \left(\frac{\alpha}{R_2} \right)^\alpha \right]^{\frac{1}{2-\alpha}} \\ &\iff \frac{A_1^2}{R_1^\alpha} > \frac{A_2^2}{R_2^\alpha}\end{aligned}\tag{2.31}$$

We observe an, at least, inverse quadratic relationship between ability and risk aversion. Intuitively, if manager 1's ability doubles, manager 2's risk aversion would have to decrease by a factor of $2^{2/\alpha}$ in order to have the same probability of winning as before the change. Hence, the probability of a given manager winning the tournament is much more sensitive to changes in ability than it is to changes in risk aversion.

2.6 Conclusion

In this paper, we introduce a tournament CEO selection model in which managers vying for the position of CEO are heterogeneous over two dimensions: ability and risk aversion. We set up the managers' utility maximization problem and show that the objective function will be the expected value of the transformation of the random variable that describes the manager's cash flow. Furthermore, for utility functions belonging to the HARA class, we can transform the objective function into a linear function of the moment generating function of the given random variable. We then solve for two such functions of the HARA class, CARA (exponential) and CRRA (iso-elastic) utility functions.

Solving the model for a representative manager's optimal choice of mean and volatility reveals the different roles that risk aversion and ability play in determining the cash flow outcomes. We find closed form expressions for mean and variance pairs for both CARA and CRRA preferences. These are both increasing in ability and decreasing in risk aversion. Specifically, we show a concave relationship between the two traits. Though the results differ quantitatively according to the specific utility function used, some interesting common patterns arise from the analysis. Namely, the concave relationship between A_i and R_i holds some important implications for CEO selection. For example, this relationship implies that, at some levels of A_i and R_i , ability and risk aversion play equally important roles in determining the winner of the CEO tournament. Hence, a manager's level of risk aversion may mask her true ability level. That is, the outcomes being used to evaluate a manager for the CEO position are not dictated only by her ability, but are also a result of her risk aversion. In conjunction with previous studies in the literature that have found that high CEO risk aversion generates negative returns for the firm, this may indicate that the firm's inability to differentiate talent from risk aversion may, at times, lead to the suboptimal choice of a CEO.

Furthermore, the relative importance of risk aversion with respect to ability increases as the level of technology increases, and vice versa. This implies that some industries, or types of firms, will be more sensitive to the CEO's actual level of risk aversion and ability than others. That is, the ability to differentiate between the two dimensions of manager quality will be more important for some firms than for others. This suggests that some firms should be willing to invest more resources into determining ability and risk aversion of CEO candidates.

Finally, we also analyze a more realistic decision based solely on observed cash flow and we are able to derive an expression for the probability of a given manager winning the tournament and being promoted to CEO. This expression is a function of her characteristics, as well as of the characteristics of the rest of the candidate pool. We derive the natural results of her probability of winning is increasing with respect to ability and decreasing with respect to risk aversion. This approach seems to be a good starting point for future research: a more sophisticated model allowing for a manager to estimate or have prior probabilities of the other candidates would potentially modify her optimal project choice, especially if the managers have heterogeneous preferences over being promoted.

REFERENCES

- [1] Dye, R. A. 1984. "The trouble with tournaments." *Economic Inquiry* 22 (1): 147–149.
- [2] Geweke, J. 2001. "A Note on Some Limitations of CRRA Utility." *Economic Letters* 71 (3): 341–45.
- [3] Goel, A. M. and A. V. Thakor. 2007. "Overconfidence, CEO Selection, Corporate Governance." *Journal of Finance* 63 (6): 2737–2784.
- [4] Green, J. R. and N. L. Stokey. 1983. "A Comparison of Tournaments and Contracts." *Journal of Political Economy* 91 (3): 349–364.
- [5] Hart, O. 1995. "Firms, Contracts, and Financial Structure." *Oxford: Oxford University Press*.
- [6] Holmstrom, B. and P. Milgrom. 1987. "Aggregation and Linearity in Provision of Intertemporal Incentives." *Econometrica* 55: 303–328.
- [7] Jensen, M. and K. J. Murphy. 1990. "Performance Pay and Top Management Incentives." *Journal of Political Economy* 98: 225–264.
- [8] Lazear, E. P. and S. Rosen. 1981 "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89 (5): 841–864.
- [9] Lazear, E. P. and P. Oyer. 2007. "Personnel Economics." In *Handbook of Organizational Economics*. Edited by R. Gibbons and J. Roberts. Forthcoming 2012. Princeton: Princeton University Press.
- [10] Nalebuff, B. J. and J. E. Stiglitz. 1983. "Prizes and Incentives: Towards a General Theory of Compensation and Competition." *The Bell Journal of Economics* 14 (1): 21–43.

- [11] Pestien, V. and W. Sudderth. 1988. "Continuous-Time Casino Problems." *Mathematics of Operations Research* 13 (2): 364–376.
- [12] Ramakrishnan, R. T. S. and A. V. Thakor. 1991. "Cooperation versus Competition in Agency." *Journal of Law, Economics, & Organization* 7 (2): 248–283.
- [13] Rosen, S. 1986. "Prizes and Incentives in Elimination Tournaments." *American Economic Review* 76 (4): 701–715.
- [14] Tsoulouhas, T., C. R. Knoeber and A. Agrawal. 2007. "Contests to Become CEO: Incentives, Selection and Handicaps." *Economic Theory* 30 (2) : 195–221.

CHAPTER 3

DIVIDEND POLICIES AND FIRM SURVIVAL

3.1 Introduction

In the business world, dividend policies are the result of an enviable situation for the firm: given a certain level of profits, should it reinvest this windfall back into the firm, or should it pay out a certain amount to its shareholders? Black (1976) described this question, and the lack of consensus for a satisfying answer, as the “dividend puzzle”. He noted that “the harder we look at the dividend picture, the more it seems like a puzzle, with pieces that just don’t fit together”. Although this remark was meant to be descriptive at the time, it surprisingly still holds today – despite the enormous amount of research performed in this area, there does not yet exist an empirical or theoretical consensus on optimal dividend policies. The contribution of this paper to this growing literature is to apply the stochastic dynamic programming techniques developed for solving optimal consumption and investment problems, especially in Sethi, Taksar and Presman (1992, 1995) and Presman and Sethi (1996a, b), to the context of a CEO who determines the dividend payouts of the firm. By utilizing the CEO tournament results from Chapter 2, this paper studies how the relationship between the CEO’s ability level and risk aversion affects dividend policies and firm survival.

As argued previously, I abstract from contract theory by assuming that the firm offers the new CEO a linear contract. The new CEO essentially becomes a partner of the firm and receives a certain number of shares.¹ These shares pay

¹This practice often takes place in large firms, consulting groups, etc. There are two rationales

out dividends and it is assumed that the CEO cannot sell her shareholdings.² The selected CEO maximizes the utility of total expected and discounted dividends. The CEO's utility function is defined by her degree of risk aversion, and her level of ability will define how much revenue she is able to generate through her investment decisions. By modifying the Sethi, Taksar and Presman (1992) model, their results can be reconciled with the model presented in this paper. Following this analysis, the CEO is then assumed to have CRRA preferences, which allows for the derivation of optimal dividend and investment policies of the CEO. I find that CEOs with either high levels of risk aversion and ability or low levels of these same traits will pay out a low dividend yield, whereas CEOs with medium ones will pay out a higher dividend yield.

As for the analysis of firm survival, most of the literature concerns itself, naturally, in modeling and testing the profit maximizing hypothesis. The hypothesis that firms are profit maximizers is obviously pervasive throughout economics, and the most common argument given for it is a market selection rationale.³ Nevertheless, it should be noted that the idea that firms are always profit maximizing, or that long run behavior of the economy is equivalent to an economy composed solely of profit maximizing firms, has not gone unquestioned. Specifically, Dutta and Radner (1999), Koopmans (1957) and Blume and Easley (2002), suggest otherwise. However, most of the literature assumes, either when modeling a manager or CEO that controls the firm or modeling the firm's preferences directly, that the firm is risk neutral or that the CEO's incentives are aligned with those of the shareholders. None of these papers allow for the CEO's possible risk aversion to play a role. For instance, Dutta and

for this: retention of the executives and incentives alignment.

²Empirically, it has been observed that equity compensation is primarily composed of restricted stocks that can be sold only after a certain time.

³See Friedman (1953), Alchian (1950) and Nelson & Winter (1982), among others.

Radner (1999) define the entrepreneur or manager as risk neutral and investigate the profit maximizing assumption. In their continuous-time setting model, they find that profit maximizing firms fail for sure in finite time, while some non profit-maximizing strategies for which firms have a positive probability of surviving do exist. This paper modifies this assumption of risk neutrality by assuming a risk averse CEO who controls the policies of the firm. In this context, I find that the probability of survival, while converging to either zero or one when risk aversion tends to either zero or infinity, is a non-monotonic function that will depend on the dividend policy implemented by the CEO.

Finally, Appendix C contains an alternate model of dividend policies and long-run evolution of the firm following a similar approach used in Dutta and Radner (1999). Using this modeling technique, I find similar results in that the CEO dividend policy is concave and non-monotonic with respect to risk aversion.

3.2 General Model

The model presented in this paper is based on the model of optimal investment and consumption in continuous time developed in Sethi, Taksar and Presman (1992, 1995), which is an extension of Karatzas, Lehoczky, Sethi and Shreve (1986), which in turn is a generalization of Lehoczky, Sethi and Shreve (1983). I reinterpret the consumption variable of their models as withdrawals to be paid out as dividends, and proceed to analyze the optimal decisions and long run dynamics of the firm. This analysis is performed with respect to one of the main results in Chapter 2: in the context of a CEO tournament, where the firm does

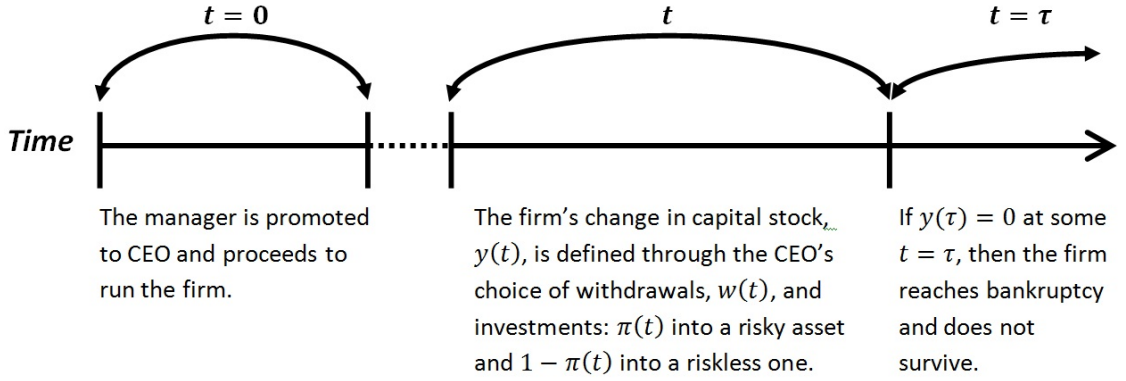


Figure 3.1: Heuristic Description of the Model

not have full information over the characteristics of the managers in the candidate pool, a concave trade-off between ability and risk aversion is expected. In other words, it is found that managers of high ability and high risk aversion may be equally as likely to be promoted as low ability and low risk aversion managers. However, these combinations of traits may drive distinct dividend and investment policies, which in turn could affect the firm's probability of survival. The setup of the model is as follows: the CEO begins her time at the firm with an initial amount of capital y . She then continues to run the firm either over an infinite time horizon or until it goes bankrupt. Figure 3.1 provides a heuristic description of the model in discrete time.

3.2.1 The CEO's Decision Problem

The CEO is faced with the decision problem of investing the firm's capital stock in two different assets: a riskless asset, with a rate of return $r > 0$; and a risky asset, with an expected rate of return $\alpha(1 + A) > r$. The risky asset's expected rate of return, as shown below, is an increasing function of the CEO's level of

ability. This setup is deceptively general: Karatzas, Lehoczky, Sethi and Shreve (1986) show that by making use of the mutual fund theorem (Merton, 1971), a model with $n > 2$ risky investments is equivalent to a model with a single risky asset. The riskless asset is modeled as:

$$\frac{dP_0(t)}{P_0(t)} = rdt \quad (3.1)$$

where P_0 is the price of the riskless asset.

The risky investment is described, as in Merton (1971) and Black and Scholes (1973), through what is usually called the "geometric Brownian motion" hypothesis. However, the usual formulation is modified to allow for the rate of return of the risky asset to be a function of the CEO's ability, A . This formulation differs slightly from Chapter 2: in this context, low ability cannot diminish the usual rate of return of the risky asset. In other words, a CEO with zero ability can expect the minimum expected rate of return α , as shown above, while CEOs with positive levels of ability will be able to invest in an asset with a higher expected rate of return. The reason for this assumption will become clear when defining the CEO's choice variables. Given this modification, I define the risky asset in the following manner:

$$\frac{dP(t)}{P(t)} = \alpha(A + 1)dt + \sigma dB(t) \quad (3.2)$$

P is the price of the risky asset. $B(\cdot)$ is a standard Brownian motion, where standard implies a constant drift of 0 and volatility of 1, while $\sigma > 0$. The stochastic process that satisfies the Itô stochastic differential equation in (3.2) follows a geometric Brownian motion.

The CEO's preferences are defined by an instantaneous utility function which is continuously differentiable, strictly concave, strictly increasing and sublinear:

$$U : \mathbb{R}_+ \rightarrow \mathbb{R}, \quad U'(w) > 0, \quad U''(w) < 0, \quad \forall w > 0, \quad \text{and} \quad U'(\infty) = \lim_{w \rightarrow \infty} U'(w) = 0$$

The CEO's intertemporal preferences are also defined by her discount rate, β , which reflects how she will value utility in future periods. Although a specific utility function is not immediately specified for the general model, when deriving the main results I assume that the CEO is characterized by her relative coefficient of risk aversion, R .

The firm will offer the CEO a contract that is linear with respect to dividends paid to the shareholders, $\eta + \lambda w(t, P)$. In other words, when the manager is promoted to CEO, she will become a shareholder.⁴ Again, as in Chapter 2, the assumptions of $\eta = 0$ and $\lambda = 1$ are made. When running the firm, the CEO chooses an optimal withdrawal policy, $w(t, P)$, $\forall t \geq 0$, as well as an optimal investment policy, $\pi(t, P)$, $\forall t \geq 0$. The withdrawals $w(t, P)$ will be distributed amongst the shareholders of the firm. In the model $\pi(t, P)$ is unconstrained, which implies unlimited borrowing and short-selling. Note that, due to the assumption that $\alpha > r$ for all ability levels, short-selling will never occur. However, the lack of a credit constraint is an obvious limitation of the model, but one that can be motivated by assuming that the firm is big enough to have significant access to capital markets. Another implication of the lack of a credit constraint is that, when defining ability through the rate of return of the risky asset, if the CEO's level of ability is such that it diminishes the baseline rate of re-

⁴See Section 2.2 of Chapter 2 for a discussion on linear contracts. Also, see Sung (2005) for a discussion on the optimality of linear contracts in a continuous-time adverse selection model.

turn, and the CEO has perfect information over her low ability, then the model allows for generating profits as if the manager had a positive level of ability by short-selling against the risky asset. Since this unintended mechanism betrays the main intention of assuming heterogeneous levels of ability – increased performance by higher skilled individuals – it is restricted to be only return-augmenting by assuming the ability parameter to be additively separable from the base expected return of the risky asset.

As shown in Karatzas, Lehoczky, Sethi and Shreve (1986) and based on Harrison and Kreps (1979) and Harrison and Pliska (1981), the firm's capital stock, for any given set of policies $(w(t, P), \pi(t, P))$, will follow the Itô stochastic differential equation:

$$dy(t, P) = \begin{cases} (\alpha * (A + 1) - r)\pi(t, P)y(t, P)dt + (ry(t, P) - w(t, P))dt \\ + \sigma y(t, P)\pi(t, P)dB(t) \end{cases} \quad (3.3)$$

$$y(0, P) = y \quad (3.4)$$

One of the main benefits of using the Sethi, Taksar and Presman (1992) model is that their subsistence constraint, defined as s , such that $w(t, P) \geq s \geq 0$ can be advantageously reinterpreted. Instead of representing subsistence consumption, in this model s will serve as the minimum level of dividends that the firm needs to pay out to its shareholders.⁵ To allow for this, two more assumptions over the utility function are needed: $U(w) = -\infty$ for $w < s$, and $U(s) = \lim_{w \rightarrow s^+} U(w)$. These assumptions merely indicate that the utility function is right-continuous up to $w = s$, with a jump occurring as soon as dividends fall below s , where utility drops to $-\infty$. In addition, it should be noted that the

⁵See DeAngelo, DeAngelo and Skinner (1992) for motivation of sticky dividend policies.

case of $s = 0$ is studied by Karatzas, Lehoczky, Sethi and Shreve (1986) and its solution requires a slightly different than the $s > 0$ case due to technical matters. However, since this paper will not delve into the derivations, this will not affect the concurrent examination of both cases and the subsequent economic interpretations.

In order to study the long term effects of the CEO's optimal policies, the model also allows for the firm to run out of working capital (i.e., it runs into bankruptcy). The firm will run into bankruptcy at time T_0 if its capital stock reaches zero at that time. Properly defined,

$$T_0 \equiv \{\sup t \geq 0 : y(\tau) = 0, \quad \forall \tau \in [0, t]\} \quad (3.5)$$

If the firm reaches bankruptcy, the CEO will receive a payment P at the time of bankruptcy T_0 . Another useful characteristic of this model is that this payment can be interpreted as a severance (or penalty) package: if the CEO runs the firm into bankruptcy, she is fired at time T_0 and receives a payment which provides her utility P . I will use the natural interpretation of $P > 0$ as a severance while $P < 0$ will be referred to as a penalty. Additionally, although I will not make further use of this remark in the rest of the paper, it should be noted that, for a large enough amount of initial capital $y > s/r$, since the CEO can assure a total payout of s in perpetuity by always investing in the riskless asset, any payment P which provides less than the expected utility of withdrawing s in perpetuity, i.e., $P \geq U(s)/\beta$, can be labeled as a reward and the reverse, $P < U(s)/\beta$, as a penalty.

Given these assumptions, the CEO's optimization problem can now be properly set up as:

$$V_{w(\cdot), \pi(\cdot)}(y, P) \equiv \max_{w(\cdot), \pi(\cdot)} E \left[\int_0^{T_0} e^{-\beta t} U(w(t, P)) dt + P e^{-\beta T_0} \right] \quad (3.6)$$

Under some weak conditions on the utility function and on the policies $w(t, P)$ and $\pi(t, P)$ to ensure the existence of a solution to Equation (3.2), the value function can be defined as:⁶

$$V(y, P) = \sup_{w(\cdot), \pi(\cdot)} V_{w(\cdot), \pi(\cdot)}(y, P) \quad (3.7)$$

As one might expect, the parameter P plays an intricate role in solving this dynamic programming problem. Before examining the characterizations of the optimal policies, it is useful to define the following values:

$$\bar{P} = \frac{U(\infty)}{\beta} \quad \text{and} \quad \tilde{P} = \begin{cases} -\infty, & \text{if } s > 0 \\ U(0)/\beta, & \text{if } s = 0 \end{cases} \quad (3.8)$$

These two values represent important cut-off values in the CEO's decision. Specifically, for a large enough reward, $P \geq \bar{P}$, the CEO's optimal strategy is to withdraw as much as possible instantaneously in order to reach bankruptcy.⁷ On the other hand, once the reward is small enough, $P < \tilde{P}$, or, in other words, the penalty is large enough, the CEO's optimal strategy will be the same as if $P = \tilde{P}$. Because of these two results, the model can be fully characterized by focusing on $P \in [\tilde{P}, \bar{P})$.

Finally, it can be shown that the value function $V(y, P)$ satisfies the Hamilton-Jacobi-Bellman equation:

⁶See Karatzas, Lehoczky, Sethi and Shreve (1986), Sethi, Taksar and Presman (1992) and Presman and Sethi (1996a) for the derivations and proofs of the base model.

⁷Since the model is in continuous time, she will not be able to withdraw the entire capital stock instantly and therefore no optimal dividend policy exists.

$$\left\{ \begin{array}{l} \beta V(y, P) = \max_{\pi, w \geq s} \left\{ \begin{array}{l} (\alpha(A+1) - r)\pi y V'(y, P) \\ + (ry - w)V'(y, P) \\ + \frac{1}{2}(\sigma\pi y)^2 V''(y, P) + U(w) \end{array} \right\}, \quad \text{if } y > 0 \text{ and } P \in [\tilde{P}, \bar{P}) \\ V(0, P) = P, \quad \text{if } y = 0 \text{ and } P \in [\tilde{P}, \bar{P}) \\ \lim_{y \rightarrow \frac{s}{r}} V(y, \tilde{P}) = \frac{U(s)}{\beta}, \quad \text{if } y \rightarrow s/r \text{ and } P = \tilde{P} \end{array} \right. \quad (3.9)$$

The solution to the H-J-B equation, the optimal dividend and investment policies chosen by the CEO, are stated in the next subsection.

3.3 Optimal Dividend and Investment Policies

For a given capital stock $y(t, P)$ at time t , where $y(0) = y$, the optimal dividend and investment policies $(w^*(t, P), \pi^*(t, P))$, $\forall t \geq 0$, that the CEO will pursue are:

$$w^*(t, P) = \begin{cases} (U')^{-1} V'(y, P), & \text{if } 0 < V'(y, P) \leq U'(s) \\ s, & \text{if } V'(y, P) \geq U'(s) \end{cases} \quad (3.10)$$

$$\pi^*(t, P)y(t, P) = -\frac{\alpha(A+1) - r}{\sigma^2} \frac{V'(y, P)}{V''(y, p)} \quad (3.11)$$

Following the analysis and summary detailed in Sethi (1998), a synopsis of my results including the results derived in the previous section for $P < \tilde{P}$ and $P > \bar{P}$ can be presented as follows:

- I. For $P > \bar{P}$ and any capital stock $y > 0$, the CEO will try to pay out dividends in such a way that bankrupts the firm as quickly as possible.
- II. For $P \in (\tilde{P}, \bar{P})$: The firm will survive with a positive probability q (and, therefore, will go bankrupt with a positive probability $1 - q$). This is driven by the optimal capital stock process, $y^*(t, P)$. The probability of survival will be equal to zero if $\beta \geq r + \gamma$, where $\gamma = \frac{(\alpha(A + 1) - r)^2}{2\sigma^2}$. In other words, the probability of survival will be zero if the CEO's discount rate is high enough. Since γ is a function of the CEO-specific risk premium $\alpha(A + 1) - r$, for ease of interpretation, γ can be thought of as the risk premium.
- (a) For all $P \in (\tilde{P}, \bar{P})$, if the Inada condition $\lim_{w \rightarrow s} U'(w) = \infty$ is satisfied, then the CEO will pay out linear withdrawals with respect to capital stock, and $w(t, P) > s, \forall t$. This last result is driven by the extreme aversion the CEO will have to paying out only the minimum level of dividends, s . As noted earlier, this condition can be better interpreted if dividends are thought of as sticky: once the precedent of paying s is established, it is very difficult to reduce dividends below this level.
- (b) For $0 < U'(s) < \infty$, withdrawals will be bounded away from s only if P is big enough: $\forall s > 0, \exists P^*(s)$ such that $\forall P \in (P^*(s), \tilde{P}), w(t, P) > s$. For the case of $P \leq P^*(s)$, $\exists \bar{y}(t, P)$ such that, for capital stocks $y(t, P) \in [0, \bar{y}(t, P)]$, only the minimum dividend is paid out, $w(t, P) = s$, while $w(t, P) > s$ when $y(t, P) > \bar{y}(t, P)$. Hence, dividends are flat at s for low enough capital stocks, and linear in capital stock when the capital stock is large enough.

III. For $P = \tilde{P}$:

- (a) The penalty for bankruptcy is so great that, for a large enough capital stock of ($y > s/r$), the firm can and will avoid bankruptcy with certainty ($q = 1$), and, if $U(s) = -\infty$, the CEO will manage the firm in such a way that the capital stock is bounded below by s/r . The dividend policy will be defined by $U'(s)$. If $U'(s) = \infty$, the optimal dividend policy will pay out $w(t, P) \geq s$, with equality occurring when $y = s/r$. If $U'(s) < \infty$, which implies $U(s) > -\infty$, then for $\tilde{y}(s) = \lim_{P \rightarrow \tilde{P}^+} \bar{y}(t, P)$, withdrawals will be either greater than the minimum, $w(t, P) > s$, if the capital stock is large enough ($y(t, P) > \tilde{y}(s)$), or will be at the minimum, $w(t, P) = s$, if the capital stock is small enough ($0 < y(t, P) < \tilde{y}(s)$).
- (b) If the capital stock is exactly $y = s/r$ and $U(s) > -\infty$, the CEO will pay out $w(t) = s$ in perpetuity by always investing in the riskless asset. In this scenario, there is no bankruptcy ($q = 1$).
- (c) For any capital stock $y \in (0, s/r)$, or for $y = s/r$ and $U(s) = -\infty$, any dividend policy the CEO wishes to implement will generate $-\infty$ utility. In these scenarios, any dividend policy is optimal and the probability of survival will be any value $q \in [0, 1)$, depending on the given withdrawal policy. Intuitively, it should be noted that this case should not be economically feasible since the firm is almost doomed to fail – a firm with an initial amount of capital which is less than what is needed to generate the minimum level of dividends through the riskless asset is essentially a inordinately risky gamble. If the recently promoted manager had the choice between either accepting or rejecting the promotion, she would surely reject it in this scenario.

IV. For $P < \tilde{P}$ and any capital stock y , the CEO will run the firm as if $P = \tilde{P}$.

It is interesting to note that, since β is the CEO's individual discount rate and r is the market rate of interest, if $\beta = r$ (the CEO values the future in the same manner as the market), then the probability of survival is zero only when $\gamma \leq 0$. Since γ is nonnegative by definition, this means that the firm will fail only in a scenario where the severance is too large and the CEO's ability is equal to zero. In other words, if the set of CEOs with ability equal to zero is a null set, the firm will have a positive probability of survival in all scenarios. Moreover, it will only have a positive probability of bankruptcy if $P > \tilde{P}$. Hence, when $\beta = r$, any failure of the firm can be traced back to CEO impatience, driven by severance penalties that are too small, or a large discount rate (or both). It should be noted, however, that, $\beta \neq r$ is a perfectly reasonable assumption, and could be driven by reasons such as pressure from the current board of directors. For this reason, stating that any failure of the firm must be due to CEO impatience is potentially too strong of an interpretation of this result.

In the following section I use results (I) through (IV) to analyze how different CEO types from a given iso-mean curve (as derived in Chapter 2) will behave and their probability of survival, specifically in the case of CRRA utility.

3.4 CRRA Preferences

To generate the main results of this paper, the CEO's instantaneous preferences are assumed to be defined by a Constant Relative Risk Aversion (CRRA) utility function. Specifically:

$$U(w) = \begin{cases} \frac{w^{1-R}}{1-R}, & \text{if } w(t) \geq s \text{ and } R > 0, R \neq 1 \\ \ln(w), & \text{if } w \geq s \text{ and } R = 1 \\ -\infty, & \text{if } w < s \end{cases} \quad (3.12)$$

where, again, the withdrawal constraint with respect to s is implemented by assuming an infinite disutility over withdrawals less than s .

Given the assumption of a CRRA utility function for the CEO, the following results can be derived:

- I. For all finite values of P , $P \leq \bar{P} = \infty$. Hence, driving the firm into bankruptcy as soon as possible is never optimal for the CEO.
- II. Since $s > 0$, $U'(s) = \frac{1}{s^R} > 0$, $\forall R > 0$. Therefore, for all values of $P \in (\tilde{P}, \infty)$, the dividend payout will be at least the minimum level s . Furthermore, $\forall s > 0, \exists P(s^*)$ such that for $P > P^*(s)$, dividends will be bounded away from s , $w(t) > s$; and, for $P \leq P^*(s)$, there will exist a capital level $\bar{y}(t, P)$, such that the CEO will only pay the minimum level s if the current capital stock level is below $\bar{y}(t, P)$ and will pay dividends $w(t, P) > s$ if the capital stock exceeds this given capital level.
- III. If $P = \tilde{P} = -\infty$, the CEO, through her optimal policies, will ensure that the firm never goes bankrupt for all capital stocks above s/r . In these scenarios, there will exist a capital level $\bar{y}(t, P)$ such that the CEO will pay dividends above s if the capital stock is larger than $\bar{y}(t, P)$ and will pay exactly s for capital stocks less than $\bar{y}(t, P)$. Lastly, if the capital stock is below s/r , bankruptcy cannot be avoided with certainty and any policy will be optimal, with differing probabilities of survival between zero and one.

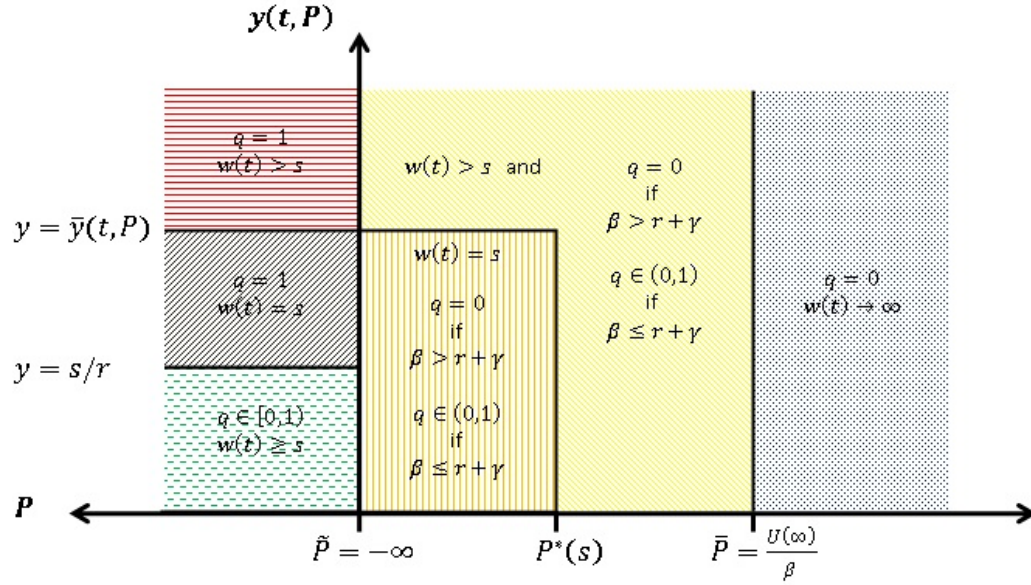


Figure 3.2: Dividend Policies and Probability of Survival as a Function of Capital Stock and Severance.

For ease of understanding, these results are depicted in Figure 3.2. Furthermore, for the appropriate cases in which the CEO pays out dividends above s , the dividend and investment policies will be:

$$w^*(t, P) = \left[\frac{\beta - r(1 - R)}{R} - \frac{\gamma(1 - R)}{R^2} \right] y(t, P) \quad (3.13)$$

$$\pi^*(t, P) = \frac{\alpha(A + 1) - r}{\sigma^2 R} \quad (3.14)$$

Intuitively, the CEO pays out a fraction of the firm's capital stock in dividends in each time period, and it also invests a fixed fraction of the capital stock in the risky asset.

3.4.1 Dividend Policies

For this subsection, one of the main results from Chapter 2 is imported: it was found that, in a CEO tournament where the firm does not have a mechanism to completely disassociate managers' ability and level of risk aversion, there exists a trade-off between the two traits which is described by a concave relationship. This implies that managers with, say, a combination of high ability and high risk aversion, are as likely to be promoted to CEO as managers with a comparable combination of low ability and a low level of risk aversion. It should be noted, however, that only the qualitative result is implemented, and not the specific relationships that were found in Chapter 2. This is done for two reasons: (1) the relationships in Chapter 2 share the concavity property over the specifications examined, but the functional forms themselves are sensitive to these specifications; and, (2) the definition of ability in this chapter, although similar in that it is return-augmenting, is not mathematically equivalent to the one in Chapter 2 due to the modifications made to accommodate the lack of credit constraints of the current model. Hence, I assume that the concave relationship is simply defined by the function $A = g(R)$, where $g'(R) > 0$ and $g''(R) < 0$.

Notice that the dividend policy is a function of γ , which in turn now is a function of R . If γ and $A = g(R)$ is substituted into (3.13), the following expression can be derived:

$$\begin{aligned}
 w^*(t, P) &= \left[\frac{\beta - r(1 - R)}{R} - \frac{\gamma(1 - R)}{R^2} \right] y(t, P) \\
 &= \left[\frac{\beta - r(1 - R)}{R} - \frac{(\alpha(A + 1) - r)^2}{2\sigma^2} \frac{1 - R}{R^2} \right] y(t, P) \\
 &= \left[\frac{\beta - r(1 - R)}{R} - \frac{(\alpha(g(R) + 1) - r)^2}{2\sigma^2} \frac{1 - R}{R^2} \right] y(t, P) \tag{3.15}
 \end{aligned}$$

From this point onwards, the analysis is focused on the *dividend yield*, which shall be referred to as θ , as opposed to total dividends. It is simply the fraction of the capital stock which is paid to the shareholders, and mathematically it is the linear coefficient in Equation (3.15):

$$\theta = \frac{\beta - r(1 - R)}{R} - \frac{(\alpha(g(R) + 1) - r)^2}{2\sigma^2} \frac{1 - R}{R^2} \quad (3.16)$$

First, I find that, for all functions $g(R)$, the dividend yield is a concave function for all values of R . Furthermore, it will have a maximum for all functions $g(R)$ with degree lower or equal to $1/2$. From this result I hypothesize that the trade-off between ability must be large enough, for if it is too small, at the limit when risk aversion goes to infinity, ability will increase at a fast enough rate that expected returns will outpace the increased risk aversion and the dividend yield will diverge to infinity. Mathematically:

$$\lim_{R \rightarrow \infty} \theta = r + \frac{\alpha^2}{2\sigma^2} * \lim_{R \rightarrow \infty} \frac{g(R)^2}{R} \quad (3.17)$$

Hence, assuming that $g(R)$ is of degree $1/2$ or lower, managers on the same iso-mean curve, i.e., who have the same likelihood of being promoted to CEO per Chapter 2's results, will have different dividend policies. The managers with high ability and, by extension, high risk aversion, as well as the managers with low ability but low risk aversion, will have lower dividend yields than managers with a combination of medium levels of ability and risk aversion. Figure 3.3, which was plotted using the values $\beta = 0.5$, $r = 1$, $\alpha = 1.4$, $\sigma = 1$ and the function $g(R) = \sqrt{R}$, illustrates this result.

However, by further interpreting these results, it can be observed that for high levels of ability and their corresponding higher levels of risk aversion, the

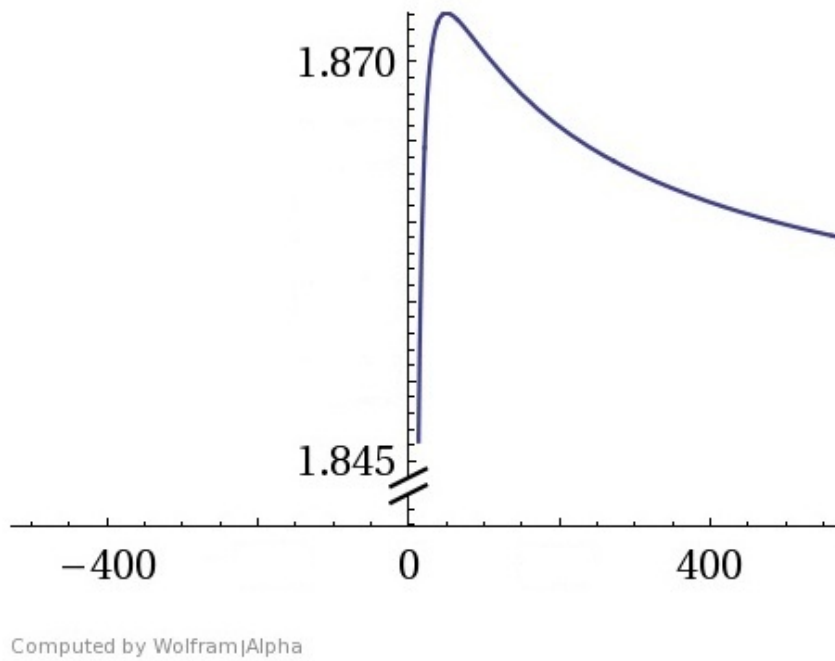


Figure 3.3: Optimal Dividend Yields as a Function of Risk Aversion for Likely CEOs.

total dividends paid out to the shareholders does not necessarily decline. Since total dividends is equal to the dividend yield times the capital stock, the change in capital stock as ability increases needs to be taken into account as well. To address this, notice how the capital stock equation changes: for a given level of capital, an infinitesimal change for a change in risk aversion will result in a change in the expected rate of return, which in turn affects the fraction of capital which is invested in the risky asset, as well as a change in the dividend yield. When risk aversion tends to infinity, the fraction invested in the risky asset converges to zero, while, at the same time, dividends will necessarily be s . Hence, at the limit, the CEO will run the firm as if it had the smallest amount of capital possible to survive (s/r). Analogously, when risk aversion tends to zero, the fraction of capital invested in the risky asset tends to infinity (the CEO would borrow an infinite amount of money), but withdrawals would also only be s . It

is obvious that these scenarios would not be optimal for the shareholders of the firm – they would much prefer a CEO with a more balanced portfolio choice.

3.5 A Note on Firm Survival

As noted in the introduction, though the model presented in this paper is based on the optimal consumption and investment literature, the main inspirations were the ideas of market selection and firm survival in Dutta and Radner (1999) and Blume and Easley (2002). Conveniently, another characteristic of adapting the Sethi, Taksar and Presman (1992) model is that the probability of survival (or, equivalently, the probability of bankruptcy) of the firm can also be analyzed. Presman and Sethi (1996b) provide a closed-form solution, as a function of the first and second derivatives of the value function, of the probability of survival:

$$P(T_0 = \infty) = \begin{cases} 1 - \left(\frac{V'(y, P)}{V'(s/r, P)} \right)^{\frac{r+\gamma-\beta}{\gamma}} & \text{if } r + \gamma - \beta > 0 \\ 0, & \text{if } r + \gamma - \beta \leq 0 \end{cases} \quad (3.18)$$

Again, the relevant comparison is to contrast potential CEOs who belong to the same iso-mean curve, or are equally as likely to be promoted to CEO. Hence, two different CEOs with coefficients of relative risk aversion R_1 and R_2 , and corresponding levels of ability $A_1 = g(R_1)$ and $A_2 = g(R_2)$ are assumed, where, without loss of generality, $R_1 > R_2$, which implies that $A_1 > A_2$. Both CEOs are assumed to be running a firm with the same amount of initial capital y , and will receive the same severance P if the firm falls into bankruptcy. Since V and γ are functions of both ability and risk aversion, when comparing the

CEOs probabilities of survival, the CEO with R_1 will have a larger probability of survival if:

$$1 - \left(\frac{V'_1(y, P)}{V'_1(s/r, P)} \right)^{\frac{r+\gamma_1-\beta}{\gamma_1}} > 1 - \left(\frac{V'_2(y, P)}{V'_2(s/r, P)} \right)^{\frac{r+\gamma_2-\beta}{\gamma_2}} \quad (3.19)$$

Since the value function, for $R \neq 1$, is right continuous at s/r and $V(s/r, P) = U(s)/\beta$, it can be easily shown that the value function for CRRA utility will be:

$$V(y, P) = \frac{y^{1-R}}{1-R} \left[\frac{\beta - r(1-R)}{R} - \frac{\gamma(1-R)}{R^2} \right]^{-R} = \frac{y^{1-R}}{1-R} \theta^{-R},$$

which allows Equation (3.19) to be simplified as:

$$\left[\beta \left(\frac{s}{r} \frac{1}{y\theta_2} \right)^{R_2} \right]^{\frac{r+\gamma_2-\beta}{\gamma_2}} > \left[\beta \left(\frac{s}{r} \frac{1}{y\theta_1} \right)^{R_1} \right]^{\frac{r+\gamma_1-\beta}{\gamma_1}} \quad (3.20)$$

Equation (3.20) demonstrates that the probability of survival with respect to R is a function of θ , the optimal dividend yield. Since θ , as shown in the previous subsection, is not monotonic with respect to R , neither is the probability of survival. Hence, which CEO will have the larger likelihood of surviving cannot be algebraically determined. It can be noted, however, that the following intuitive results indeed hold: as $R \rightarrow \infty$, the probability of survival of the firm is 1, and as $R \rightarrow 0$, the probability of survival is 0.

3.6 Conclusion

In this paper, I provide an application of the optimal consumption and investment models from the operations research literature to model the effects of CEO

characteristics on optimal dividend policies and long run dynamics of the firm. The main reason for adopting a modified version of the Sethi, Taksar and Presman (1992, 1995) model is to take advantage of the closed form solutions they were able to derive for very general formulations of investment-consumption models. Assuming managers are heterogeneous with respect to risk aversion and levels of ability, the optimal dividend and investment policies of the CEO can be fully characterized and shown to be functions of her ability and risk aversion. I import one of the main results of Chapter 2: there exists a concave relationship between ability and risk aversion in terms of the likelihood of a given manager being promoted to CEO. Although this result is only implemented qualitatively, when assuming that the CEO is characterized by a CRRA utility function, the optimal dividend policy of the CEO can be shown to vary non-monotonically with respect to risk aversion. Specifically, CEOs with combinations of either high ability and risk aversion, or low ability and risk aversion, will implement dividend policies with a lower yield than CEOs with a more balanced combination of ability and risk aversion.

Furthermore, by implementing the results of Presman and Sethi (1996b), the probability of survival of the firm given a type of CEO can be defined and explored in detail. I find that the probability of survival is a function of the optimal dividend yield of a given CEO, and this function is not monotonic. Specifically, there exists a range of risk aversion coefficients for which the risk-averse behavior of the CEO counteracts a potentially higher dividend yield, and the resulting change in the probability of survival is ambiguous. However, the following natural results are verified: as risk aversion (and, by extension, ability) converges to either zero or infinity, the probability of survival of the firm converges to either zero or one, respectively.

Finally, in Appendix C, I include a continuous-time model á la Dutta and Radner (1999) which also uses the risk aversion-solution set solution set previously found in Chapter 2, and in this context the optimal dividend policy can again be characterized in terms of levels of risk aversion and ability. Assuming CRRA preferences and defining cash flow as a geometric Brownian motion, the same result as in the paper is found: CEOs with levels of risk aversion and ability either both high or both low will pay out a low dividend yield, whereas CEOs with medium ones will pay out a high dividend yield.

REFERENCES

- [1] Alchian, A. 1950. "Uncertainty, Evolution and Economic Theory." *Journal of Political Economy* 58 (3): 211–221.
- [2] Black, F. 1976. "The Dividend Puzzle." *Journal of Portfolio Management* 2 (2): 5–8.
- [3] Black, F. and Scholes, M. J. 1972. "The Valuation of Option Contracts and a Test for Market Efficiency." *Journal of Finance* 27 (2): 399–418.
- [4] Black, F. and Scholes, M. J. 1973. "The Policy of Options and Corporate Liabilities." *Journal of Political Economy* 81 (3): 637–654.
- [5] Blume, L. and D. Easley. 1992. "Evolution and Market Behavior." *Journal of Economic Theory* 58 (1): 9–40.
- [6] Blume, L. and D. Easley. 2002. "Optimality and Natural Selection in Markets." *Journal of Economic Theory* 107 (1): 95–135.
- [7] DeAngelo, H., L. DeAngelo and D. J. Skinner. 1992. "Dividends and Losses." *Journal of Finance* 47 (5): 1837–1863.
- [8] Dutta, P. K. and R. Radner. 1999. "Profit Maximization and the Market Selection Hypothesis." *Review of Economic Studies* 66 (4): 769–798.
- [9] Friedman, M. 1953. *Essays in Positive Economics*. Chicago: University of Chicago Press.
- [10] Harrison, J. M. and D. M. Kreps. 1979. "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory* 20 (3): 381–408.

- [11] Harrison, J. M. and S. R. Pliska. 1981. "Martingales and Stochastic Integrals in the Theory of Continuous Trading." *Stochastic Processes and their Applications* 11 (3): 215–260.
- [12] Karatzas, I., J. P. Lehoczky, S. P. Sethi and S. E. Shreve. 1986. "Explicit Solution of a General Consumption/Investment Problem." *Mathematics of Operations Research* 11 (2): 261–294.
- [13] Karatzas, I. and S. E. Shreve. 1988. *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.
- [14] Koopmans, T. C. 1957. *Three Essays on the State of Economic Science*. New York: McGraw-Hill.
- [15] Lehoczky, J. P., S. E. Shreve and S. P. Sethi. 1983. "Optimal Consumption and Investment Policies Allowing Consumption Constraints and Bankruptcy." *Mathematics of Operations Research* 8 (4): 613–636.
- [16] Merton, R. C. 1969. "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case." *Review of Economics and Statistics* 51 (3): 247–257.
- [17] Merton, R. C. 1971. "Optimum Consumption and Portfolio Rules in a Continuous Time Model." *Journal of Economic Theory* 3 (4): 373–413.
- [18] Merton, R. C. 1973. "Erratum." *Journal of Economic Theory* 6 (2): 213–214.
- [19] Merton, R. C. 1975. "Theory of Finance from the Perspective of Continuous Time." *Journal of Financial and Quantitative Analysis* 10 (4): 659–674.
- [20] Miller, M. H. and F. Modigliani. 1961. "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business* 34 (4): 411–33.

- [21] Miller, M. H. and K. Rock. 1985. "Dividend Policy under Asymmetric Information." *Journal of Finance* 40 (4): 1031–51.
- [22] Nelson, R. and S. G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press.
- [23] Presman, E. and S. P. Sethi. 1996a. "Consumption Behavior in Investment/Consumption Problems." In *Optimal Consumption and Investment with Bankruptcy*, edited by S. P. Sethi, 185–206. Boston: Kluwer.
- [24] Presman, E. and S. P. Sethi. 1996b. "Distribution of Bankruptcy Time in a Consumption/Portfolio Problem." *Journal of Economic Dynamics and Control*, 20 (1–3): 471–477.
- [25] Presman, E. and S. P. Sethi. 1996c. "Risk-Aversion Behavior in Consumption/Investment Problems with Subsistence Consumption." In *Optimal Consumption and Investment with Bankruptcy*, edited by S. P. Sethi, 155–184. Boston: Kluwer.
- [26] Radner, R. 1998. "Profit Maximization with Bankruptcy and Variable Scale." *Journal of Economic Dynamics and Control* 22 (6): 849–867.
- [27] Radner, R. and L. Shepp. 1996. "Risk vs. Profit Potential: A Model for Corporate Strategy." *Journal of Economic Dynamics and Control* 20 (8): 1373–1393.
- [28] Sethi, S. P. 1998. "Optimal Consumption-Investment Decisions Allowing for Bankruptcy: A Survey." In *Worldwide Asset and Liability Modeling*, edited by W. T. Ziemba and J. M. Mulvey, 397–426. Cambridge, UK: Cambridge University Press.

- [29] Sethi, S. P., M. Taksar and E. Presman. 1992. "Explicit Solution of a General Consumption/Portfolio Problem with Subsistence Consumption and Bankruptcy." *Journal of Economic Dynamics and Control* 16 (3): 747-768.
- [30] Sethi, S. P., M. Taksar and E. Presman. 1995. "Erratum: Explicit Solution of a General Consumption/Portfolio Problem with Subsistence Consumption and Bankruptcy." *Journal of Economic Dynamics and Control* 19 (4): 1297-1298.
- [31] Sung, J. 2005. "Optimal Contracts Under Adverse Selection and Moral Hazard: A Continuous-Time Approach." *Review of Financial Studies* 18 (3): 1021–1073.

APPENDIX A
APPENDIX TO CHAPTER 1

A.1 Adverse Selection Design à la Macho-Stadler and Pérez-Castrillo

A.1.1 Firm Behavior

Firms will maximize profits given their characteristics and what contract they are being offered: the regulator observes $\tilde{\beta}$ and $\tilde{\theta}$ and offers a menu of contracts (p_f, R)

The firms' maximization problem will be:

$$\max_{(\tilde{\beta}, \tilde{\theta}) \in [\underline{\beta}, \bar{\beta}] \times [\underline{\theta}, \bar{\theta}]} \left\{ p_a A \left(R \left(\tilde{\beta}, \tilde{\theta} \right) \right) + p_f \left(\tilde{\beta}, \tilde{\theta} \right) * F \left(R \left(\tilde{\beta}, \tilde{\theta} \right), \tilde{\beta} \right) \right\} \quad (\text{A.1})$$

For tractability purposes, we will further simplify this problem to

$$\max_{\tilde{\gamma} \in [\underline{\gamma}, \bar{\gamma}]} \left\{ p_a \gamma \left(\bar{R} - v \left(R \left(\tilde{\gamma} \right) \right) \right) + T \left(\tilde{\gamma} \right) \right\} \quad (\text{A.2})$$

where $v(\cdot)$ is an increasing and convex function, which represents the efficiency of the land to be fallowed/sequester carbon/used to produce carbon offsets, and $\gamma \equiv \frac{\theta}{\beta}$, is the collapsed, relative benefit of producing agriculture as opposed to offsets.

Note that for the contract menu to make sense, the regulator must be sure that a type- γ agent will be interested in accepting the contract $(R(\gamma), T(\gamma))$,

instead of untruthfully declaring herself a type $-\tilde{\gamma}$ in order to sign the contract $(R(\tilde{\gamma}), T(\tilde{\gamma}))$. This is what is referred to as the incentive compatibility condition (*IC*) and we can write it as:

$$\begin{aligned} p_a \gamma \left(\bar{R} - v(R(\gamma)) \right) + T(\gamma) &\geq p_a \gamma \left(\bar{R} - v(R(\tilde{\gamma})) \right) + T(\tilde{\gamma}) \\ \iff T(\gamma) - p_a \gamma v(R(\gamma)) &\geq T(\tilde{\gamma}) - p_a \gamma v(R(\tilde{\gamma})) \end{aligned} \quad (\text{A.3})$$

Also, in order for the firm to accept the contract, it must receive at least the reservation utility. This is known as satisfying the individual rationality condition (*IR*). In our case, this is derived from allocating the entire plot of land towards agriculture:

$$\begin{aligned} p_a \gamma \left(\bar{R} - v(R(\gamma)) \right) + T(\gamma) &\geq \max \{ p_a \gamma \bar{R}, 0 \} \\ \iff \begin{aligned} T(\gamma) - p_a \gamma v(R(\gamma)) &\geq 0, & \text{if } \gamma \geq 0 \\ T(\gamma) - p_a \gamma v(R(\gamma)) &\geq -p_a \gamma \bar{R}, & \text{if } \gamma < 0 \end{aligned} \end{aligned} \quad (\text{A.4})$$

For both scenarios, we can see that the right hand side of the inequalities are constant. We will solve for both cases separately.

A.1.2 Regulator

Thanks to the revelation principle, we can concentrate on “point contracts”, given that the corresponding incentive compatible constraints hold. The optimization problem then becomes:

$$\max_{R(\gamma), T(\gamma)} \int_{\underline{\gamma}}^{\bar{\gamma}} \left[\Pi(R(\gamma)) - T(\gamma) \right] f(\gamma) d\gamma$$

subject to

$$\begin{cases} (IR1) & T(\gamma) - p_a \gamma v(R(\gamma)) \geq 0 & \text{for all } \gamma \in [0, \bar{\gamma}] \\ (IR2) & T(\gamma) - p_a \gamma v(R(\gamma)) \geq -p_a \gamma \bar{R} & \text{for all } \gamma \in [\underline{\gamma}, 0] \\ (IC) & T(\gamma) - p_a \gamma v(R(\gamma)) \geq T(\tilde{\gamma}) - p_a \gamma v(R(\tilde{\gamma})) & \text{for all } \gamma, \tilde{\gamma} \in [\underline{\gamma}, \bar{\gamma}] \end{cases}$$

As noted in Macho-Stadler and Pérez-Castrillo, this problem has some special characteristics. The following properties are important for the solution:

- I. The *Spence–Mirrlees (single crossing) condition*: $\frac{\partial}{\partial \gamma} \left(\frac{\partial \pi / \partial R}{\partial \pi / \partial T} \right) = \frac{\partial}{\partial \gamma} \left(\frac{-p_a \gamma v'(R)}{1} \right) = -p_a v'(R) < 0$.
- II. We can divide the problem into two stages:
 - (a) First, for each $R(\gamma)$ that the regulator wants the firm to use, she must find the transfer $T(\gamma)$ that ensures that each agent type will effectively choose that contract $(R(\gamma), T(\gamma))$.
 - (b) Second, between all the functions $R(\gamma)$ for which a transfer function $T(\gamma)$ can be found, the regulator must choose that which maximizes expected welfare. Since the function satisfies the S-M condition, the only way in which the regulator can implement a particular land requirement $R(\gamma)$ is if the function satisfies the condition $\frac{dR}{d\gamma} \leq 0$. Hence, land required for offsets is decreasing in γ (increasing in relative efficiency for offsets).
- III. Only the individual rationality constraint of the least efficient type is needed. If this one is satisfied, all individual rationality constraints for more efficient types are automatically satisfied:

$$T(\gamma) - p_a \gamma v(R(\gamma)) \geq T(\bar{\gamma}) - p_a \gamma v(R(\bar{\gamma})) \geq T(\bar{\gamma}) - p_a \bar{\gamma} v(R(\bar{\gamma})) \geq 0$$

IV. The incentive compatibility constraint can be transformed in the following manner:

$$\pi(\gamma) \equiv \max_{\gamma^o} \left\{ T(\gamma^o) - p_a \gamma v(R(\gamma^o)) \right\} = T(\gamma) - p_a \gamma v(R(\gamma))$$

By the envelope theorem, we know that $\frac{d\pi(\gamma)}{d\gamma} = -p_a v(R(\gamma))$. Hence, $\pi(\gamma)$ can be written as:

$$\pi(\gamma) = \pi(\bar{\gamma}) + \int_{\bar{\gamma}}^{\gamma} \frac{d\pi(x)}{dx} dx = \pi(\bar{\gamma}) + \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx \quad (\text{A.5})$$

Given that the participation constraint for $\bar{\gamma}$ must be satisfied, we have

$$\pi(\gamma) = T(\gamma) - p_a \gamma v(R(\gamma)) = \bar{U} + \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx \quad (\text{A.6})$$

$$\text{where } \bar{U} = \begin{cases} 0, & \text{if } \gamma \geq 0 \\ -p_a \gamma \bar{R}, & \text{if } \gamma < 0 \end{cases}$$

The proof required to show $R(\gamma)$ is implementable if and only if it is not increasing in γ is omitted. Since our last expression implies that the participation constraint for $\bar{\gamma}$ is satisfied, we can rewrite the regulators problem as:

$$\max_{R(\gamma), T(\gamma)} \int_{\gamma}^{\bar{\gamma}} \left[\Pi(R(\gamma)) - T(\gamma) \right] f(\gamma) d\gamma \quad (\text{A.7})$$

$$\text{subject to } \begin{cases} \frac{dR(\gamma)}{d\gamma} \leq 0 \\ T(\gamma) = p_a \gamma v(R(\gamma)) + \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx + \bar{U} \end{cases} \quad (\text{A.8})$$

$$\text{where, again, } \bar{U} = \begin{cases} 0, & \text{if } \gamma \geq 0 \\ -p_a \gamma \bar{R}, & \text{if } \gamma < 0 \end{cases}$$

By substituting $T(\gamma)$ into the objective function, we get

$$\max_{R(\gamma)} \int_{\underline{\gamma}}^{\bar{\gamma}} \left[\Pi(R(\gamma)) - p_a \gamma v(R(\gamma)) - \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx - \bar{U} \right] f(\gamma) d\gamma \quad (\text{A.9})$$

$$\text{subject to } \frac{dR(\gamma)}{d\gamma} \leq 0 \quad (\text{A.10})$$

Assuming the constraint is satisfied, we integrate by parts to arrive at:

$$\begin{aligned} & \int_{\underline{\gamma}}^{\bar{\gamma}} \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx f(\gamma) d\gamma \\ &= \left[F(\gamma) \int_{\gamma}^{\bar{\gamma}} p_a v(R(x)) dx \right]_{\underline{\gamma}}^{\bar{\gamma}} - \int_{\underline{\gamma}}^{\bar{\gamma}} p_a v(R(x)) F(\gamma) d\gamma \\ &= - \int_{\underline{\gamma}}^{\bar{\gamma}} p_a v(R(x)) F(\gamma) d\gamma = - \int_{\underline{\gamma}}^{\bar{\gamma}} p_a v(R(x)) \frac{F(\gamma)}{f(\gamma)} f(\gamma) d\gamma \end{aligned}$$

Substituting in the regulators problem gives us

$$\left\{ \max_{R(\gamma)} \int_{\underline{\gamma}}^{\bar{\gamma}} \left[\Pi(R(\gamma)) - p_a v(R(\gamma)) \left(\gamma + \frac{F(\gamma)}{f(\gamma)} \right) - \bar{U} \right] f(\gamma) d\gamma \right. \quad (\text{A.11})$$

The FOC of this problem is:

$$\Pi'(R(\gamma)) - p_a v'(R(\gamma)) \left(\gamma + \frac{F(\gamma)}{f(\gamma)} \right) = 0 \implies v'(R(\gamma)) = \frac{\Pi'(R(\gamma))}{\gamma + \frac{F(\gamma)}{f(\gamma)}} \quad (\text{A.12})$$

Furthermore, if we define $\Pi(R(\gamma)) = p_a \gamma (\bar{R} - v(R(\gamma))) + p_e R(\gamma)$ we have

$$v'(R(\gamma)) = \frac{p_e - p_a \gamma v'(R(\gamma))}{\gamma + \frac{F(\gamma)}{f(\gamma)}} \quad (\text{A.13})$$

For a uniform distribution, we have

$$v'(R(\gamma)) = \frac{p_e - p_a \gamma v'(R(\gamma))}{2\gamma - \underline{\gamma}} \quad (\text{A.14})$$

Compare to the result with full information, which should be:

$$v'(R(\gamma)) = \frac{p_e - p_a \gamma v'(R(\gamma))}{\gamma} \quad (\text{A.15})$$

From this we can determine that under asymmetric information, while there is no “distortion at the top”, i.e., for $\underline{\gamma}$ the results are the same. However, for every other value of γ we have reduced efficiency in that the regulator demands less land for offsets under adverse selection.

Also, note that putting some structure over $v(\cdot)$ would allow us to determine $R(\gamma)$, which in turn would allow us to determine $T(\gamma)$ (through the incentive compatibility condition). This fully characterizes the contract.

Finally, $\frac{dR(\gamma)}{d\gamma} \leq 0$ holds if and only if

$$\left[\Pi''(R(\gamma)) - p_a v''(R(\gamma)) \left(\gamma + \frac{F(\gamma)}{f(\gamma)} \right) \right] \frac{dR(\gamma)}{d\gamma} - p_a v'(R(\gamma)) \left(1 + \frac{d}{d\gamma} \frac{F(\gamma)}{f(\gamma)} \right) = 0 \quad (\text{A.16})$$

Note that, since the SOC of the regulators problem is satisfied:

$$\Pi''(R(\gamma)) - p_a v''(R(\gamma)) \left(\gamma + \frac{F(\gamma)}{f(\gamma)} \right) < 0$$

we have

$$\begin{aligned} \frac{dR(\gamma)}{d\gamma} \leq 0 &\iff 1 + \frac{d}{d\gamma} \frac{F(\gamma)}{f(\gamma)} \geq 0 \\ &\iff \frac{d}{d\gamma} \frac{F(\gamma)}{f(\gamma)} \geq -1 \end{aligned}$$

which is a property satisfied by several distributions (including the uniform).

Let us see what this contract looks like by using functional forms. Specifically, if we assume the land production function $v(\cdot)$ to be $v(R) = \sqrt{R}$, we have that $v'(R) = \frac{1}{2\sqrt{R}}$, and

$$\frac{1}{2\sqrt{R(\gamma)}} = \frac{p_e - p_a \gamma \frac{1}{2\sqrt{R(\gamma)}}}{2\gamma - \underline{\gamma}}$$

$$R(\gamma) = \left[\frac{(2 + p_a) \gamma - \underline{\gamma}}{2p_e} \right]^2 \quad (\text{A.17})$$

Now,

$$T(\gamma) = p_a \gamma \ln R(\gamma) + p_a \int_{\gamma}^{\bar{\gamma}} v(R(x)) dx + \bar{U} \quad (\text{A.18})$$

Integrating by parts:

$$\begin{aligned} \int_{\gamma}^{\bar{\gamma}} v(R(x)) dx &= \int_{\gamma}^{\bar{\gamma}} \ln \left(\frac{(2 + p_a) x - \underline{\gamma}}{p_e} \right) dx \\ &= \left[\left(x - \frac{\underline{\gamma}}{2 + p_a} \right) \ln \left(\frac{(2 + p_a) x - \underline{\gamma}}{p_e} \right) - x \right]_{\gamma}^{\bar{\gamma}} \\ &= \left(\bar{\gamma} - \frac{\underline{\gamma}}{2 + p_a} \right) \ln \left(\frac{(2 + p_a) \bar{\gamma} - \underline{\gamma}}{p_e} \right) - \bar{\gamma} - \gamma \ln R(\gamma) \\ &\quad + \left(\frac{\underline{\gamma}}{2 + p_a} \right) \ln R(\gamma) + \gamma \end{aligned} \quad (\text{A.19})$$

Substituting Equation (A.19) into Equation (A.18) results in:

$$\begin{aligned}
T(\gamma) &= p_a \gamma \sqrt{R(\gamma)} + p_a \int_{\gamma}^{\bar{\gamma}} v(R(x)) dx + \bar{U} \\
&= p_a \bar{\gamma} \left[\frac{(2 + p_a) \gamma - \underline{\gamma}}{2p_e} \right] + p_a (\bar{\gamma} - \gamma) \left[\frac{(2 + p_a) (\bar{\gamma} + \gamma) - 2\underline{\gamma}}{4p_e} \right] + \bar{U} \\
&= 2p_a \left[\frac{(2 + p_a)}{2p_e} [2\bar{\gamma}\gamma + (\bar{\gamma} - \gamma)(\bar{\gamma} + \gamma)] + \underline{\gamma}\gamma \right] + \bar{U} \tag{A.20}
\end{aligned}$$

APPENDIX B

APPENDIX TO CHAPTER 2

In this Appendix, we provide the details on how to derive the optimization solutions for the CARA/Normal and CRRA/Log-Normal specifications for the CEO Selection Model in Chapter 2.

B.1 Solution for CARA/Normal Distribution Specification

$$\max_{(\mu_i, \sigma_i)} E \left[-\frac{\exp(-R_i \cdot X)}{R_i} \right] \quad (\text{B.1})$$

$$\text{subject to} \quad \begin{cases} \mu_i \leq A_i \cdot (\sigma_i)^\alpha \\ X \sim N(\mu_i, \sigma_i^2) \end{cases} \quad (\text{B.2})$$

Let $Y_i = -R_i \cdot X$. Since Y_i is a linear transformation of X , we know that Y_i also follows a normal distribution, i.e., $Y_i \sim N(-R_i\mu_i, R_i^2\sigma_i^2)$. Then, our optimization problem is equivalent to:

$$\min_{(\mu_i, \sigma_i)} \frac{1}{R_i} E [\exp(Y)] \quad (\text{B.3})$$

$$\text{subject to} \quad \begin{cases} \mu_i \leq A_i \cdot (\sigma_i)^\alpha \\ Y_i \sim N(-R_i\mu_i, R_i^2\sigma_i^2) \end{cases} \quad (\text{B.4})$$

We also know that if Y_i is normally distributed, then $\exp(Y_i)$ follows a log-normal distribution: $\exp(Y_i) \sim \log -N(-R_i\mu_i, R_i^2\sigma_i^2)$. This allows us to further simplify our optimization problem:

$$\min_{(\mu_i, \sigma_i)} \frac{1}{R_i} \left[\exp \left(-R_i \mu_i + \frac{R_i^2 \sigma_i^2}{2} \right) \right] \quad (\text{B.5})$$

$$\text{subject to} \quad \mu_i \leq A_i \cdot (\sigma_i)^\alpha \quad (\text{B.6})$$

which is equivalent to:

$$\max_{(\mu_i, \sigma_i)} \left\{ R_i \mu_i - \frac{R_i^2 \sigma_i^2}{2} \right\} \quad (\text{B.7})$$

$$\text{subject to} \quad \mu_i \leq A_i \cdot (\sigma_i)^\alpha \quad (\text{B.8})$$

The technology constraint will bind at the optimum ($\mu_i = A_i \cdot \sigma_i^\alpha$). The objective function will be strictly concave (an inverted parabola) with respect to σ_i . Hence, the first-order condition is necessary and sufficient to characterize the solution:

$$\frac{\partial \left(R_i A_i \cdot \sigma_i^\alpha - \frac{R_i^2 \sigma_i^2}{2} \right)}{\partial (\sigma_i)} = \alpha R_i A_i \cdot \sigma_i^{\alpha-1} - R_i^2 \sigma_i = 0 \quad (\text{B.9})$$

Solving for μ_i and σ_i :

$$\sigma_i = \left(\frac{\alpha A_i}{R_i} \right)^{\frac{1}{2-\alpha}} \quad (\text{B.10})$$

$$\mu_i = (A_i)^{\frac{2}{2-\alpha}} \left(\frac{\alpha}{R_i} \right)^{\frac{\alpha}{2-\alpha}} \quad (\text{B.11})$$

B.2 Solution for CRRA/Log-Normal Distribution Specification

$$\max_{(\mu_i, \sigma_i)} E \left[\frac{X^{1-R_i}}{1-R_i} \right] \quad (\text{B.12})$$

$$\text{subject to} \quad \begin{cases} \mu_i \leq A_i \cdot (\sigma_i)^\alpha \\ \log(X) \sim N(\mu_z, \sigma_z^2) \end{cases} \quad (\text{B.13})$$

where μ_i and σ_i are the mean and variance of the log-normally distributed random variable X . We can manipulate the expected utility function to get a more manageable optimization problem:

$$\begin{aligned}
E \left[\frac{X^{1-R_i}}{1-R_i} \right] &= \frac{1}{1-R_i} E \left[\exp \log \left(X^{1-R_i} \right) \right] \\
&= \frac{1}{1-R_i} E \left[\exp \{ (1-R_i) \log (X) \} \right] \\
&= \frac{1}{1-R_i} M_z (1-R_i) \\
&= \frac{1}{1-R_i} \exp \left[\mu_z (1-R_i) + \frac{\sigma_z^2 (1-R_i)^2}{2} \right] \\
&= \frac{1}{1-R_i} \left\{ \exp \left[\mu_z + \frac{\sigma_z^2 (1-R_i)}{2} \right] \right\}^{1-R_i}
\end{aligned}$$

where $M_z(\cdot)$ is the moment generating function of z .

Since the manager is choosing between mean-variance pairs for cash flow (μ_i, σ_i) , we need to express μ_z and σ_z in terms of (μ_i, σ_i) , or equivalently, in terms of (μ_i, σ_i^2) . From $z \sim N(\mu_z, \sigma_z^2)$, $X = \exp(z) \sim \log - N(\mu_z, \sigma_z^2)$ we have:

$$\begin{aligned}
\mu_i &= \exp \left(\mu_z + \frac{\sigma_z^2}{2} \right) \\
\sigma_i^2 &= \left[\exp(\sigma_z^2) - 1 \right] \exp(2\mu_z + \sigma_z^2)
\end{aligned}$$

Solving this system for μ_z and σ_z^2 results in:

$$\begin{aligned}
\mu_z &= \log \left[\frac{\mu_i^2}{(\sigma_i^2 + \mu_i^2)^{\frac{1}{2}}} \right] \\
\sigma_z^2 &= \log \left[\frac{\sigma_i^2 + \mu_i^2}{\mu_i^2} \right]
\end{aligned}$$

which implies

$$\begin{aligned}
E \left[\frac{X^{1-R_i}}{1-R_i} \right] &= \frac{1}{1-R_i} \left\{ \exp \left[\log \left[\frac{\mu_i^2}{(\sigma_i^2 + \mu_i^2)^{\frac{1}{2}}} \right] + \left(\frac{1-R_i}{2} \right) \log \left[\frac{\sigma_i^2 + \mu_i^2}{\mu_i^2} \right] \right] \right\}^{1-R_i} \\
&= \frac{1}{1-R_i} \left\{ \exp \left[(1+R_i) \log(\mu_i) - \frac{R_i}{2} \log(\sigma_i^2 + \mu_i^2) \right] \right\}^{1-R_i} \\
&= \frac{1}{1-R_i} \left\{ \frac{\mu_i^{1+R_i}}{(\sigma_i^2 + \mu_i^2)^{\frac{R_i}{2}}} \right\}^{1-R_i}
\end{aligned}$$

Now, after a monotone increasing transformation on this objective function, we can solve an equivalent maximization problem:

$$\max_{(\mu_i, \sigma_i)} \left\{ \frac{\mu_i^{\frac{2}{R_i}+2}}{\sigma_i^2 + \mu_i^2} \right\} \quad (\text{B.14})$$

$$\text{subject to} \quad \mu_i \leq A_i \cdot \sigma_i^\alpha \quad (\text{B.15})$$

Since the technology constraint binds at the optimum ($\mu_i = A_i \cdot \sigma_i^\alpha$), we have:

$$\max_{\mu_i} \left\{ \frac{\mu_i^{\frac{2}{R_i}}}{\frac{\mu_i^{\frac{2}{\alpha}(1-\alpha)}}{(A_i)^{\frac{2}{\alpha}}} + 1} \right\} \quad (\text{B.16})$$

If α and R_i satisfy the additional constraint $R_i > \frac{\alpha}{1-\alpha}$, the objective function will be strictly concave over $\mu_i \geq 0$. Hence, the first-order condition is necessary and sufficient to characterize the solution:

$$\frac{\partial}{\partial \mu_i} \left[\mu_i^{\frac{2}{R_i}} \left(\frac{\mu_i^{\frac{2}{\alpha}(1-\alpha)}}{(A_i)^{\frac{2}{\alpha}}} + 1 \right)^{-1} \right] =$$

$$= \frac{2\mu_i^{\frac{2}{R_i}-1}}{R_i \left[\frac{\mu_i^{\frac{2}{\alpha}(1-\alpha)}}{(A_i)^{\frac{2}{\alpha}}} + 1 \right]} - \frac{\mu_i^{\frac{2}{R_i}}}{\left[\frac{\mu_i^{\frac{2}{\alpha}(1-\alpha)}}{(A_i)^{\frac{2}{\alpha}}} + 1 \right]^2} \left(\frac{2(1-\alpha)}{\alpha A_i^{\frac{2}{\alpha}}} \mu_i^{\frac{2}{\alpha}-3} \right) = 0 \quad (\text{B.17})$$

From this condition, the closed-form expression for μ_i can be found:

$$\begin{aligned} \Leftrightarrow \mu_i^{\frac{2}{\alpha}(1-\alpha)} &= \frac{\alpha \left[A_i^{\frac{2}{\alpha}} + \mu_i^{\frac{2}{\alpha}(1-\alpha)} \right]}{(1-\alpha) R_i} \\ \Leftrightarrow \mu_i^{\frac{2}{\alpha}(1-\alpha)} \left[\frac{R_i}{\alpha} - R_i - 1 \right] &= A_i^{\frac{2}{\alpha}} \\ \Leftrightarrow \mu_i &= \left[\frac{A_i^{\frac{2}{\alpha}}}{\frac{R_i}{\alpha} - R_i - 1} \right]^{\frac{\alpha}{2(1-\alpha)}} \end{aligned} \quad (\text{B.18})$$

Finally, by plugging in Equation (B.18) into the binding technology constraint (Equation B.15), the solution for σ_i can be derived:

$$\sigma_i = \left[\frac{A_i}{\left(\frac{R_i}{\alpha} - R_i - 1 \right)^{\frac{1}{2}}} \right]^{\frac{1}{1-\alpha}} \quad (\text{B.19})$$

APPENDIX C
APPENDIX TO CHAPTER 3

C.1 General Model à la Dutta and Radner (1999)

Firm behavior is modeled in continuous time to study its optimal dividend policy and implications on survival. The earnings or cash flow process of the firm is modeled as a controlled diffusion which the CEO controls through a mean-variance framework and through which her level of ability will also play a role. Specifically, the CEO chooses $z(t) = [m(Y(t), A), s(Y(t), A)] \in Z(Y(t), A)$, where $m(t)$ and $s(t)$ are the drift and the volatility of a given project, A is her personal level of ability and $Y(t)$ is her capital stock at time t (the firm starts at period $t = 0$ with an initial capital stock $Y(0) = y > 0$). With this definition it is implicitly assumed that the optimal controls are stationary. This means that $z(t) = z(t'), \forall t, t' \in \mathbb{R}_{++}, t \neq t' \iff Y(t) = Y(t')$.

The cash flows stochastic process of the firm is defined as follows: Let $F(\cdot)$ be a filtration generated by a standard Brownian motion $B(\cdot)$. $X(t)$ is the solution, adapted to $F(\cdot)$, of the stochastic differential equation:

$$dX(t) = \begin{cases} m(X(t), t) dt + s^{\frac{1}{2}}(X(t), t) dB(t), & 0 < t < T \\ 0, & t > T \end{cases} \quad (\text{C.1})$$

where T is the “stopping” time or the first t at which accumulated cash flow reaches zero, i.e., $X(T) = 0$. T is defined as the moment when the firm becomes bankrupt. Note that T does not necessarily have to be finite. If cash flow never reaches zero, then $T = \infty$. Also, the starting point of the cash flow is defined as the initial capital stock y : $X(0) = y$.

The capital stock flow equation at time t is given by $dY(t) = dX(t) - w(t)dt$, where $w(t)$ represents instantaneous dividends paid out to investors at time t . This implies that the capital stock flow equation can be expressed as:

$$dY(t) = \begin{cases} [m(X(t), t) - w(t)] dt + s^{\frac{1}{2}}(X(t), t) dB(t), & \text{for } 0 < t < T \\ 0, & \text{for } t > T \end{cases}$$

$W(t) = \int_0^t w(u)du$ represents accumulated dividends paid out up until time t . Assuming there exists a solution that satisfies Equation C.1, it is also known that $W(t) \leq y + X(t), \forall t$.

The CEO maximizes her expected utility:

$$\max_{W(t) \leq y + X(t)} V(y) = E \left[\int_0^T e^{-rt} u(\bar{w} + \beta w_t) dt \mid Y(0) = y \right] \quad (\text{C.2})$$

$$\text{subject to } \begin{cases} dY(t) = [m(X(t), t) - w(t)]dt + s^{\frac{1}{2}}(X(t), t)dB(t) \\ W(t) = \int_0^t w(u)du \wedge [m(t), s(t)] \in Z(Y(t), A) \end{cases} \quad (\text{C.3})$$

where $r > 0$ is an exogenously given market interest rate.

C.1.1 Assumptions about Technology

Through these assumptions I model how the entrepreneur controls the cash flow process. Here is where the ability of a certain type of CEO will matter: all else equal, CEO's with higher ability will have a larger choice set than CEO's with lower ability. Hence, if every parameter in the model is fixed except ability, a CEO with higher ability will perform at least as well as a CEO with lower ability. Formally, these assumptions are:

1. Z is compact-valued and has a non-empty intersection: $\bigcap Z(Y(t), A) \neq \emptyset \iff \exists(m, s) \in Z(Y(t), A), \forall t, A$.
2. Investment is inherently risky: $\forall(m, s) \in Z(Y(t), A), s > 0, \forall Y(t) \in \mathbb{R}_+, \forall t, A$.
3. Decreasing returns with respect to capital: Define $m(y, A) \equiv \max\{m \mid (m, s) \in Z(Y(t), A)\}$. This means that, given r , $m(y + \varepsilon) < m(y) + r\varepsilon$.
4. $Z(Y(t), A)$ is monotone increasing with respect to $Y(t)$ and A : $y_1 < y_2 \implies Z(y_1, \bar{A}) \subseteq Z(y_2, \bar{A})$ and $A_1 < A_2 \implies Z(\bar{y}, A_1) \subseteq Z(\bar{y}, A_2)$.

C.2 CRRA Preferences/Geometric Brownian Motion Specification

For this section, I assume the CEO is characterized by a CRRA utility function $u(w) = \frac{w^{1-R}}{1-R}$ and a technology set where she does not directly control the cash flow process, i.e., $Z(Y(t), A) = Z = (mAY(t), sY(t)^2)$, where (m, s) is a fixed drift-volatility pair. The CEO does, however, control the cash flow process indirectly: her choice of $w(t)$ at time t will influence the capital stock at time $t + \varepsilon$.¹ Since (m, s) is fixed, the stochastic process that describes cash flow reduces to a geometric Brownian motion:

$$dX(t) = m(X(t), t) dt + s^{\frac{1}{2}}(X(t), t) dB(t) = mAX(t) dt + s^{\frac{1}{2}}X(t) dB(t)$$

¹Radner and Shepp (1996) argue that describing firm profits with the multiplicative Black and Scholes (1972) model, which is equivalent to our geometric Brownian motion specification, is inappropriate for an individual firm. However, their reasoning follows from the fact that they assume risk neutrality for the firm, which would imply constant returns to scale profits. In this model, given the risk averse nature of the CEO, this issue need not be addressed.

which is a stochastic differential equation with solution:

$$X(t) = y \exp \left[s^{\frac{1}{2}} B(t) + \left(mA - \frac{1}{2}s \right) t \right]$$

From this solution, the dynamics of the firm's capital stock can be expressed as:

$$dY(t) = mAY(t)dt + s^{1/2}Y(t)dB(t) - w(t)dt$$

Then, the CEO is faced with choosing a dividend stream that maximizes her expected utility function. Applying the same simplifications as in the previous section (namely, $\bar{w} = 0$ and $\beta = 1$), the CEO will maximize the expected utility of total discounted dividends:

$$\max_{W(t)} V(y) = E \left[\int_0^T \exp(-rt) u(w_t) dt \mid Y(0) = y \right] \quad (\text{C.4})$$

$$\text{subject to } \begin{cases} dY(t) = mAY(t)dt + s^{1/2}Y(t)dB(t) - w(t)dt \\ W(t) = \int_0^t w(u) du \end{cases} \quad (\text{C.5})$$

In order to solve this problem, the optimal dividend policy is assumed (and later proven) to be of the form: $w(Y(t)) = \theta Y(t)$. Now, a solution for θ is needed. First, the capital stock dynamics can be rewritten as:

$$\begin{aligned} dY &= mAY(t)dt + s^{1/2}Y(t)dB(t) - w(Y(t)) \\ &= mAY(t)dt + s^{1/2}Y(t)dB(t) - \theta Y(t) \\ &= (mA - \theta)Y(t)dt + s^{1/2}Y(t)dB(t) \end{aligned}$$

Since the dividend policy is proportional to firm's capital, instantaneous withdrawals can be expressed as:

$$dw = (mA - \theta)w(t)dt + s^{1/2}w(t)dB(t)$$

λ is defined as the shadow price of the initial capital stock, y :

$$\lambda = \frac{dV(y)}{dy}$$

By the FOC, it must be that:

$$\lambda = \frac{du(w)}{dw} = w^{-R}$$

Next, by applying Ito's Lemma:

$$d\lambda = -Rw^{-1-R}dw + \frac{1}{2}R(1+R)w^{-2-R}dw^2$$

Moreover, since:

$$E_t dw = (mA - \theta)w(t)dt$$

and:

$$dw^2 = sw(t)^2 dt$$

The dynamics of the shadow price can be expressed as:

$$\frac{E_t d\lambda}{dt} = R w^{-1-R}((\theta - mA) + \frac{1}{2}(1+R)s)$$

However, this expression can be transformed in the following fashion:

$$\begin{aligned} \frac{E_t d\lambda}{dt} &= \lambda(r - (mA - Rs)) \\ &= w^{-R}(r - (mA - Rs)) \end{aligned}$$

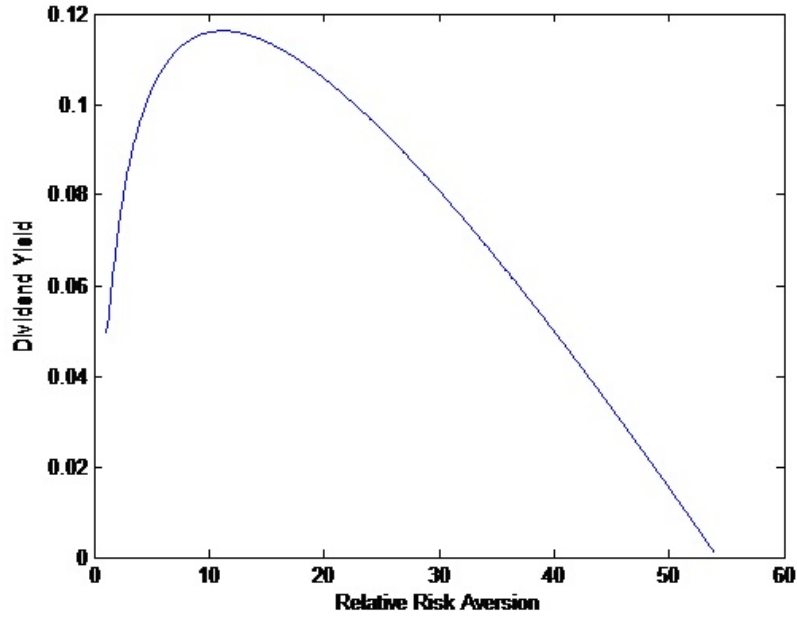


Figure C.1: Plot of the Optimal Dividend Policy

Reconciling the two previous equations, at the optimum the following condition must hold:

$$R((\theta - mA) + \frac{1}{2}(1 + R)s) = (r - (mA - Rs))$$

Finally, solving for θ :

$$\theta = mA - \frac{mA - r}{R} + \frac{1 - R}{2}s \quad (\text{C.6})$$

This fully characterizes the optimal withdrawal policy. In terms of instantaneous dividends:

$$w(t) = \left[mA - \frac{mA - r}{R} + \frac{1 - R}{2}s \right] Y(t) \quad (\text{C.7})$$

Given a specific solution set (R, A) , Figure C.1 illustrates the optimal divi-

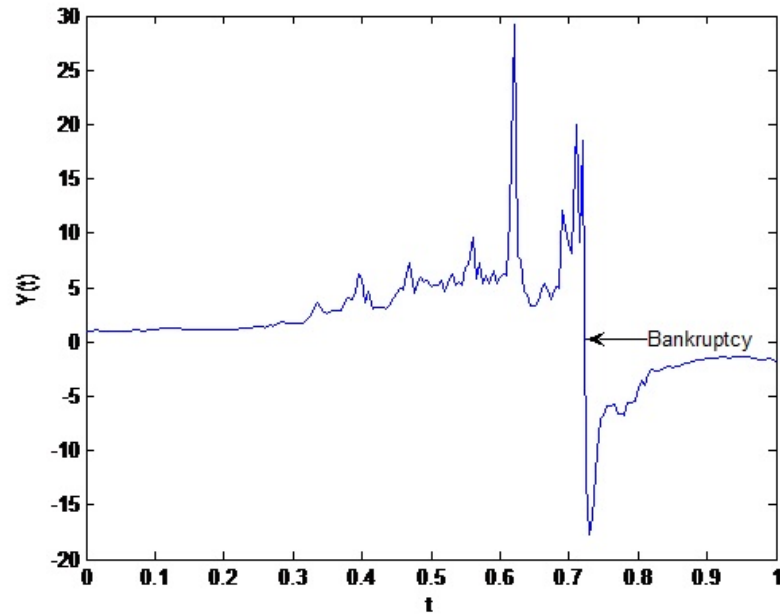


Figure C.2: Plot of a Sample Path for Capital Stock

dend policy in function of the CEO's coefficient of relative risk aversion, R . In this example the drift, m , the volatility, s , the technology parameter, α , and the interest rate, r , are equal to 0.1, 0.1, 0.5 and 0.05 respectively. The figure shows that CEOs with levels of risk aversion and ability either both high or low should payout a low dividend yield, whereas CEOs with medium ones should payout a high dividend yield.

Figure C.2 provides a sample path of the capital stock, $Y(t)$, given that firm's cash flows follow a geometric Brownian motion and that the CEO applies the optimal dividend policy. The firm goes bankrupt when its capital stock falls below zero.